

(19) 日本国特許庁 (J P)

## (12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2001-160840

(P 2 0 0 1 - 1 6 0 8 4 0 A)

(43) 公開日 平成13年6月12日 (2001.6.12)

(51) Int. Cl. <sup>7</sup>	識別記号	F I	テマコード (参考)
H04L 29/06		H04J 3/00	U
H04J 3/00		H04L 13/00	Z
H04L 12/64		11/20	A
12/28			D

審査請求 未請求 請求項の数28 O L 外国語出願 (全64頁)

(21) 出願番号	特願2000-302084 (P 2000-302084)	(71) 出願人	596077259 ルーセント テクノロジーズ インコーポ レイテッド Lucent Technologies Inc. アメリカ合衆国 07974 ニュージャージ ー、マレーヒル、マウンテン アベニュー 600-700
(22) 出願日	平成12年10月2日 (2000.10.2)	(74) 代理人	100081053 弁理士 三俣 弘文
(31) 優先権主張番号	09/411781		
(32) 優先日	平成11年10月1日 (1999.10.1)		
(33) 優先権主張国	米国 (US)		

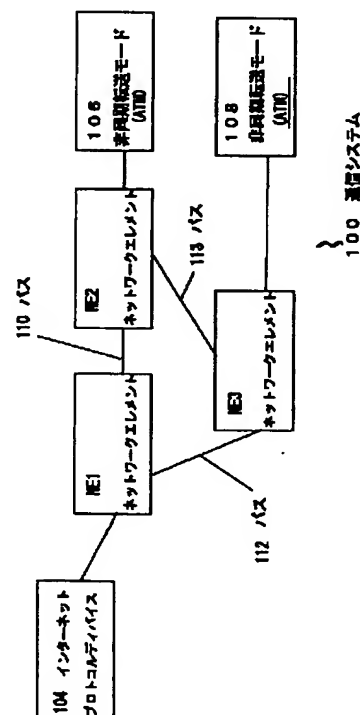
最終頁に続く

(54) 【発明の名称】 通信システムおよび通信システムにおける方法

(57) 【要約】 (修正有)

【課題】 SONET/SDHトランスポートとIPトランスポートとの間のギャップを埋めて、動作コストを最小化しかつトラフィックエンジニアリングを容易にすること。

【解決手段】 通信システムは、ネットワーク内のノード間で直接層間通信を提供するコンバージェンスプロトコルを含み、一方のネットワークノードのトランスポートレイヤは、他方のネットワークノードのサービスレイヤと直接的に通信する。サービスレイヤは、インターネットプロトコル (IP) レイヤであり、トランスポートレイヤは、SONETトランスポートレイヤである。コンバージェンスプロトコルをランさせるノードの各々は、ADMまたはデジタルクロスコネクトである。



## 【特許請求の範囲】

【請求項1】 OSI 7レイヤモデルに適合する第1のネットワークエレメントと、

OSI 7レイヤモデルに適合する第2のネットワークエレメントとを有し、

各エレメントは、コンバージェンスプロトコルを使用し、ネットワークエレメントのうちの一方のレイヤのうちの1つは、他方のネットワークエレメントのうちの異なるレイヤと直接的に通信することができることを特徴とする通信システム。

【請求項2】 前記コンバージェンスプロトコルは、一方のネットワークエレメントのうちのトランスポートレイヤが、他方のネットワークエレメントのうちのサービスレイヤと直接的に通信することを許容することを特徴とする請求項1記載のシステム。

【請求項3】 前記サービスレイヤは、インタネットプロトコルサービスレイヤであることを特徴とする請求項2記載のシステム。

【請求項4】 前記トランスポートレイヤは、同期光網(SONET)トランスポートレイヤであることを特徴とする請求項2記載のシステム。

【請求項5】 複数のパケット交換デバイスをさらに含み、前記第1および第2のネットワークエレメントが、前記パケット交換デバイスに対する回線交換クロスコネクタを提供するために接続されていることを特徴とする請求項2記載のシステム。

【請求項6】 前記パケット交換デバイスのうちの少なくとも1つは、インタネットプロトコル(IP)デバイスであることを特徴とする請求項5記載のシステム。

【請求項7】 前記パケット交換デバイスのうちの少なくとも1つは、非同期転送モード(ATM)デバイスであることを特徴とする請求項5記載のシステム。

【請求項8】 一方のネットワークエレメントは、リクエストするネットワークエレメントに帯域幅を与えることにより、リンク上の帯域幅に対するコンバージェンスプロトコルを使用して、他方のネットワークエレメントからのリクエストに応答することを特徴とする請求項4記載のシステム。

【請求項9】 他方のネットワークエレメントからの帯域幅に対する一方のネットワークエレメントからの前記リクエストは、ダイレクト通信チャネル中のインタレイヤ、インタエレメント通信によりもたらされることを特徴とする請求項8記載のシステム。

【請求項10】 1つのネットワークエレメントは、レイヤ3およびレイヤ1の動作を監視することを特徴とする請求項9記載のシステム。

【請求項11】 パケット交換ネットワークエレメントおよび回線交換ネットワークエレメントを含む通信システムにおける方法において、

(A) 1つのネットワークレイヤにおいて1つのネッ

トワークエレメントからメッセージを送信するステップと、

(B) 異なるネットワークエレメント中の異なるネットワークレイヤにおいてステップ(A)のメッセージを受信するステップとを有することを特徴とする方法。

【請求項12】 前記ステップ(A)は、

(A1) 一方のネットワークエレメントのトランスポートレイヤから、別のネットワークエレメントのサービスレイヤと直接的に通信するステップを含むことを特徴とする請求項11記載の方法。

【請求項13】 前記サービスレイヤは、インタネットプロトコルサービスレイヤであることを特徴とする請求項12記載の方法。

【請求項14】 前記トランスポートレイヤは、同期光網(SONET)トランスポートレイヤであることを特徴とする請求項12記載の方法。

【請求項15】 (C) 複数のパケット交換デバイスに対する回線交換クロスコネクタを提供するステップをさらに含むことを特徴とする請求項11記載の方法。

【請求項16】 前記ステップ(C)は、

(C1) インタネットプロトコル(IP)デバイスに対する回線交換クロスコネクタを提供するステップを含むことを特徴とする請求項15記載の方法。

【請求項17】 前記ステップ(C)は、

(C2) 非同期転送モード(ATM)デバイスに対する回線交換クロスコネクタを提供するステップを含むことを特徴とする請求項15記載の方法。

【請求項18】 (D) 1つのネットワークエレメントから追加的な帯域幅のための別のものへリクエストを送信するステップと、

(E) 追加の帯域幅を、追加の帯域幅をリクエストするネットワークエレメントに与えるステップとをさらに含むことを特徴とする請求項12記載の方法。

【請求項19】 前記ステップ(D)は、ダイレクト通信チャネル中のインタレイヤ、インタエレメント通信によりもたらされることを特徴とする請求項18記載の方法。

【請求項20】 (F) ネットワークエレメントがレイヤ3およびレイヤ1動作を監視するステップと、

(G) 前記ステップ(F)のネットワークエレメントがレイヤ3およびレイヤ1動作における変化に応答して、論理パイプの帯域幅を調節するステップとをさらに含むことを特徴とする請求項19記載の方法。

【請求項21】 ハイブリット、回線交換/パケット交換通信システムにおいて通信パスを確立する方法において、

(A) OSIスタックおよびTCP/IPスタックの両方をランニングさせることにより、SONET/SDHトランスポートと外部IPトランスポートとの間のイ

インタフェースを提供するノードを初期化するステップと、

(B) OSIスタックをランさせることにより、デジタルクロスコネクトであるSONET/SDHネットワークエレメントを初期化するステップと、

(C) 中間システム-中間システム (IS-IS) 隣接関係を使用してネットワークトポロジーを応答ディスクカバーするステップとを有することを特徴とする方法。

【請求項 22】 前記ステップ (C) は、

(D) IS-ISリンクステートプロトコルデータユニットの交換に基づいて、前記通信システム内の各ノードの物理的トポロジーを確立するステップをさらに含むことを特徴とする請求項 21 記載の方法。

【請求項 23】 (E) CAPの対間に論理パスを確立するステップをさらに含むことを特徴とする請求項 21 記載の方法。

【請求項 24】 ステップ (E) は、

(E1) IS-ISラベル交換バスメッセージの受信に応答して、CAPによりルーティングテーブルをリフレッシュするステップをさらに含むことを特徴とする請求項 23 記載の方法。

【請求項 25】 ステップ (E) は、

(E2) アドレスバインディングテーブルのCAPの確立をし、別のCAPに到達するために必要とされるパス情報を決定するステップをさらに含むことを特徴とする請求項 24 記載の方法。

【請求項 26】 (F) あて先CAPへ、コンバージェンスプロトコル帯域幅割当てメッセージを渡すステップをさらに含むことを特徴とする請求項 23 記載の方法。

【請求項 27】 (G) ステップ (F) において帯域幅割当てメッセージの受信に応答して、帯域幅をあて先CAPに割当てするステップをさらに含むことを特徴とする請求項 26 記載の方法。

【請求項 28】 (H) コンバージェンスプロトコル割当て解除メッセージの受信に応答して、あて先CAPの帯域幅割当て解除を行なうステップをさらに含むことを特徴とする請求項 27 記載の方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、通信システムに係り、特に、回線交換通信サービスとパケット交換通信サービスの統合に関する。

【0002】

【従来の技術】 高速データ伝送の使用および分化したデータサービスの需要の爆発的な伸びは、今日および将来の急増する需要を満たすために、現存する通信システム、回線交換システムとパケット交換システムの幾分うまくない結合の能力を精一杯働かせる。公衆交換電話網を使用する通常のデータ通信システムは、回線交換TD

Mトランスポートをインターネットプロトコル (IP) スイッチングのようなベストエフォートパケットスイッチングと組合せ、最大のリンク利用を強調するようにデータ伝送に影響を与える。

【0003】 通信リンクは、異なる接続方法に生じる異なる利点を有する数多くの方法で確立され得る。2つの機器間の2つ以上のチャネルの直接接続、即ちユーザに情報を交換するためのチャネルの排他的使用を提供する接続は、回線交換(circuit switched)、またはライン交換(line switched)接続と呼ばれる。回線交換は、高い信頼性のサービスを生じる技法であり、チャネルの一時的な喪失(loss)がいらいするものであり、反復的なそのような喪失が受け入れられないものである音声のような「リアルタイムの」通信アプリケーションにとって特に適している。

【0004】 回線交換は、高い信頼性のある専用線サービスにも使用される。5ESSのような電子的交換システムは、例えば時分割多重 (TDM) を使用して回線交換により、多数の電話機器を相互接続することができる。エンドユーザが適切な品質のサービスを受けることを保証するために、交換機は、典型的には、交換されるチャネルを運ぶトランクラインのアクティビティを監視しかつ周期的に検査する。

【0005】 通信誤りが生じた場合、交換機は、誤りを起こしたシステムコンポーネントを分離する、またはその正確な位置を決定するために、「ループバック」を使用することができる。故障が分離されると、データがループバックを通して故障したシステムコンポーネントを迂回できるように、または他の訂正手段を採ることができるように、システムはそれ自体を再構成(reconfigure)することができる。TDMトランスポートネットワークは、確かなレベルの性能および信頼性を提供する。

【0006】 同期光網/同期デジタル階層 (SDH/SONET) のような技術が、音声接続および専用線アプリケーションに対して、優れたジッター、ワンダー (wander) および誤り性能で、1秒あたりギガビットのレートまでスケラブルな大容量トランスポートを提供するために、トランスポートインフラストラクチャにおいて使用され得る。SONET/SDHセルフヒーリングリングは、ネットワーク故障の後数十ミリ秒内にサービスレベル回復を可能にする。

【0007】 パケット交換は、専用回線交換ラインのような通信リンクの利用を最大化するために使用され得る。パケット交換アプローチで、データは、パケット中で送信され、通信チャネルは、パケットの送信の期間だけ占有される。送信の後、チャネルは、他の機器間で転送されるパケットにより使用可能である。リンクは、最大のリンク利用を得るためにスタティスティックに多重化され、典型的には、TDMトランスポートネットワークにより専用回線上を運ばれる。

【0008】パケット交換システムは、しばしば、発信元からあて先へパケットを経路選択するために、インターネットプロトコル(IP)トランスポート法を使用する。そのようなシステムは、一般に、パケットを喪失する「ベストエフォート(best-effort)」技法を使用し、それらは一般に高い信頼性および予測可能な性能を保証する手段を欠く。スタティスティックな多重化は高いリンク利用を生じるが、IPデータネットワークにより提供されるベストエフォートサービスは、予測不可能な遅れ、ジッターおよびパケットロスに伴う。

【0009】現存するIPデータネットワークは優れた接続性を提供するが、それらは、エンドユーザに対してデータ伝送サービスを提供するためにIPデータネットワークを使用する様々なサービスプロバイダの間でネットワーク資源を制御可能に分配することができない。即ち、作業者の介入を典型的に必要とするプロビジョニング(provisioning)プロセスにより、即ち比較的頻繁に行われないプロセスにより、TDMトランスポートサービスは、各サービスプロバイダに対して固定帯域幅の通信チャネルを提供する。

【0010】パケットデータトラフィックは本来的に不規則であり、即ち大量の使用のバーストの後に比較的非アクティブな期間が続くので、TDMトランスポートの固定帯域幅の「パイプ」は、IPデータトランスポートサービスプロバイダのようなトランスポートユーザがそれらのエンドユーザからの需要の変化に対応できる柔軟性を制限する。サービスプロバイダは、非常に混雑している間は、その伝送をやめる(scale back)に違いないというエンドユーザのよい振る舞いに頼らなければならない。

【0011】よく知られた「公衆の悲劇(tragedy of the commons)」は、ユーザの大きなグループの協力的なよい振る舞いに頼ることは、長い期間において合理的なアプローチでない可能性があることを示している。これは、利用が大きい期間の間、トラフィックをやめる(scale back)という協力をしない傾向にあるストリーミングビデオのようないくつかのアプリケーションにおいて、特に明らかにされている。

【0012】また、サービスレイヤおよびトランスポートレイヤは、そのようなマルチレイヤデュアルアーキテクチャ(回線交換トランスポート/IP)通信システムにおいて分離されているので、トランスポートマネジメントは、別個の動作および保守機能に分離されている。この動作および保守機能分離は、一般に、別個の組織の調整を必要とし、チャネルのエンドツーエンド準備のプロセスを困難なタスクにする。そのような準備は、過度の量の熟練および時間を必要とし、トラフィックエンジニアリングのタスクを複雑にし、結果として、同時にトランスポートコアがあまり利用されない間、サービス品質が低下する。

【0013】回線交換システムにより提供される高いサービス品質およびパケット交換システムにより与えられる高い利用度の両方を利用する一方で、サービスプロバイダに分化したサービス保証および対応するサービスレベルアグリーメントを提供するようにネットワークインフラストラクチャサポートを提供するシステムおよび方法が、非常に望ましい。また、そのようなシステムが、劇的に増大して、バックボーンネットワークインフラストラクチャの容量を最大限に分担し、少なくとも部分的にダイナミックにトランスポートコア帯域幅を管理することにより、出現しつつあるデータアプリケーションに対する洗練されたサービス分化を提供することが大変望ましい。

【0014】

【発明が解決しようとする課題】SONET/SDHトランスポートネットワークの早い復旧、確かめられた安定性、低コストおよび低いトランスポート待ち時間を利用し、かつSONET/SDHトランスポートとIPトランスポートとの間のギャップを埋めて、動作コストを最小化しかつトラフィックエンジニアリングを容易にするシステムが強く望まれている。

【0015】

【課題を解決するための手段】本発明の原理による通信システムは、ネットワーク内のノード間の効率のよい層間通信を提供するコンバージェンスプロトコルを含む。例えば、コンバージェンスプロトコルは、1つのネットワークノードのトランスポートレイヤが別のネットワークノードのサービスレイヤと直接的に通信することを可能にする。

【0016】例示的な実施形態において、サービスレイヤは、インターネットプロトコル(IP)レイヤとして構成され、これは、非同期転送モードスイッチングを使用することが可能であり、トランスポートレイヤは、SONETトランスポートレイヤである。各ノードは、アドドロップマルチプレクサ(ADM)またはデジタルクロスコネクタのような通信エレメントであり得る。この層間通信は、レイヤ1パススルーオペレーション(layer-1 pass through operation)により得られる。

【0017】本発明の一側面において、今までスタティックであったSONET帯域幅準備が、柔軟になる。また、トランスポートとサービスとの間の区分が取り除かれる。この増大した柔軟性は、分化したサービスの保証および対応するサービスレベルアグリーメントに対するサポートを提供する。このシステムは、回線交換システムに特有の高品質のサービスおよび同時にパケット交換システムの高い利用度の特徴を示す。

【0018】本発明の原理によるシステムは、新規のコンバージェンスプロトコルによりIPトランスポートおよびSONET/SDHトランスポートを合併させる。

このアプローチにより提供される柔軟性は、リンクの内

部のトランスポート帯域幅のダイナミックな管理を可能にし、これにより、エンドツーエンド待ち時間がほとんどないバースト性のトラフィックを最小のコストで供給する。

【0019】例えば、4OC12CベースドOC48リンクのような1つのチャンネル化されたインタフェースが与えられると、通常のスタティックプロビジョニング

(provisioning)で、各支流(tributary)のターミネーションは、延長された期間、しばしば数ヶ月について固定される。しかし、本発明の原理によれば、支流の終端点(termination point)は、サービスプロバイダのランタイムのニーズに適用するようにダイナミックに変更され得る。結果として、可能性のある再構成期間は、数ヶ月から数分に、または数秒にも減少され得る

【0020】本発明の原理によれば、通信ネットワークは、以下、コンバージェンスプロトコルと呼ばれるプロトコルを使用する。コンバージェンスプロトコルは、SONET/SDHトランスポートディテールをカプセル化し、変化するサービス需要に適用するために、オンザフライ(on the fly)で内部帯域幅が管理されることを可能にする。結果として、一般に、データサービスプロバイダは、外部インターネットバックボーンとSONET/SDHトランスポートバックボーンとの間のスムーズなインタフェースにより、トランスポートネットワークインフラストラクチャの内部のディテールを気にする必要がない。

【0021】SONET/SDHトランスポートバックボーンは、その内部帯域幅を適応的に再アレンジし、自動化されたエンドツーエンドプロビジョニングを可能にすることができる。1つまたは2つ以上のネットワークエレメントが、コアアクセスポイント(CAP)として構成され、SONET/SDHトランスポートと外部IPトランスポートとの間のインタフェースを提供する。

【0022】

【発明の実施の形態】図1の概念的ブロック図において、パケット交換コンポーネントおよび回線交換コンポーネントの両方を含む通信システム100が、本発明の原理に従って動作する。トランスポートレイヤにおいて、1つまたは2つ以上のSONET/SDHネットワークエレメント(この例示的な実施形態において、NE1、NE2およびNE3)は、インターネットプロトコル(IP)デバイス104、非同期転送モード(ATM)デバイス106およびATMデバイス108のようなパケット交換デバイスに対する回線交換クロスコネク

【0023】IPデバイス104およびATMデバイス106および108は、例えば、リングトポロジーに配置され得る1つまたは2つ以上のデバイスの組合せであり得る。システム100は、本発明の説明および理解を単純化するために、限られた数のデバイスのみが示され

ている。IPデバイス104とATMデバイス106との間およびIPデバイス104とATMデバイス108との間の2つのクロスコネクバス110および112は、それぞれSONET/SDHトランスポートレイヤにおける特定の回線交換バスを提供する。

【0024】各バスは、複数のSONET/SDHネットワークエレメントを通過することができ、図1に示されていない他のバスと、これらのエレメントの各々を通る帯域幅を共有することができる。サービスプロバイダは、IPデバイス104とATMデバイス106との間およびIPデバイス104とATMデバイス108との間でデータを送信することを望む顧客による使用のために、バス110および112の各々において帯域幅の一部分をプロビジョン(provision)することができる。

【0025】以下説明されるレイヤは、OSIレイヤと呼ばれ、例えば、Ming-Chwan Chow, "Understanding SONET/SDH" ANDAN Publisher, Holmdel, NJ pages 2-31 through 2-32により説明されている。層間通信がエレメント内通信に制限されている通常のシステムと異なり、本発明による通信システムは、ネットワークエレメント間のレイヤ間通信を実行することができる。即ち、NE1のようなネットワークエレメントのサービスレイヤは、ネットワークエレメントNE2のトランスポートレイヤと通信することができる。

【0026】以下により詳細に説明するように、この通信バスは、本発明の原理によるネットワークが、バスにおける変化する帯域幅需要に適合するようにバス110および112を再プロビジョンすることを可能にする。これは、プロビジョニングがスタティックであり、かつマニュアルである通常のシステムと対称的である。即ち、通常のシステムにおいて、AOLのようなサービスプロバイダは、AT&Tのようなトランスポートサービスプロバイダによる再プロビジョニングをリクエストしなければならず、これは、作業者の介入を必要とする。そのような再プロビジョニングは、典型的には、1年に数回のみ行われる。

【0027】本発明の原理によれば、例えば、バス110が所定の伝送レートに対して、プロビジョンされ、NE1とNE2の間のリンクに対して需要のサージが生じる場合、ネットワークエレメントNE1は、NE1からNE2へのバス112および113に沿ってデータをNE2に供給するための十分な帯域幅をネットワークエレメントNE3に直接的にリクエストできる。この即時のプロビジョニングは、以下に詳細に説明する直接通信チャネル(DCC)における直接的な層間およびエレメント間通信により得られる。

【0028】図1のネットワークエレメントNE1、NE2およびNE3の各々は、CAPとして動作することができ、インターネットプロトコルを使用するもののようなパケット交換ネットワークに対するSONETラン

サポートを可能にする。フロントエンド集合(aggregate) ルータは、エントリCAPと統合されることができ、パケットレベルトラフィックはエントリCAPによりサービスされ得る。また、各CAPがレイヤ3およびレイヤ1動作の両方を監視する能力を有するので、各CAPは、個々の論理パイプのサイズを"on the fly" (他のノードとコンバージェンスプロトコルメッセージを交換することにより) 再アレンジする能力を有する。

【0029】結果として、本発明のシステムは、パイプ利用を最適化するために帯域幅を適応的に調節することができ、同時に各CAPの性能監視に基づいてベストエフォートデリバリティラフィックに対するサービスを改善する。これは、グローバルネットワークスタティスティックスを集める集中ネットワーク管理システムからの入力に基づいて帯域幅アレンジメントを調節することもできる。

【0030】図2の概念的ブロック図は、本発明の原理によるネットワークエレメント200のパケット交換コンポーネント202と回線交換コンポーネント204との間の相互接続を示す。パケット交換コンポーネント202は、ラインカード208および210に動作可能に接続されたインタネットプロトコルおよび非同期転送モードスイッチ206を含む。ラインカード208および210は、回線交換セルフ216の回線交換入力/出力214を通して回線交換セクション204へSONETリンク212により接続されている。

【0031】回線交換セクション202は、機構216、218および220のような複数のローカルスイッチ機構または回線交換シェルフを含み得る。機構の各々は、機構216のI/O214および216のようなI/O、ローカルスイッチコア222、シェルフコントロール224およびインタフェース226を含み得る。インタフェース226は、ローカルスイッチ機構とセントラルスイッチ機構228との間の通信を提供する。セントラルスイッチ機構228は、シェルフコントロール230、ローカル機構の各々に対するインタフェース232、およびセントラルスイッチコア234を含む。

【0032】図2のネットワークエレメント200のようなネットワークエレメントが、回線交換ネットワークとパケット交換ネットワークとの間のインタフェースとして使用され得る。ネットワークエレメントがこのように使用される場合、以下、これはコアアクセスポイントと呼ばれる。図3の概念的ブロック図に示されたネットワーク300は、エレメント306および308のIP伝送とエレメント310、312、314および316のSONETトランスポートとの間のインタフェースを提供するコアアクセスポイント302および304を含む

ソースノードNSAPアドレス:

あて先ノードNSAPアドレス:

レスポンス値:

む。即ち、ネットワークエレメント310、312、314、316、302および304の間の伝送は、SONET/SDHトランスポート伝送であり、ネットワークエレメント302、304、306、307および308間の伝送は、IP伝送である。

【0033】本発明の原理によれば、ネットワークエレメントは、以下、「コンバージェンスプロトコル"convergence protocol"」と呼ばれることになる新しいプロトコルとの組合せで、コアアクセスポイントとして動作する。コンバージェンスプロトコルは、OSIスタックをカプセル化する。コンバージェンスプロトコルデータユニットフォーマットは、図4に示されている。このデータユニットは、5バイトのヘッダおよび可変長フィールドを含む。ヘッダは、1バイトのコンバージェンスプロトコルディスクリミネータ、1バイト長インジケータ、長さ表示およびバージョン番号の間で分割された1バイト、1バイトのプロトコルデータユニットタイプインジケータ、およびチェックサムを含む。各プロトコルスタックは、それ自体のレイヤ構造を有する。

【0034】例えば、OSIにおいて、7個のレイヤ、物理、リンク、ネットワーク、トランスポート、セッション、プレゼンテーション、およびアプリケーションがある。この例示的な実施形態において、新しいプロトコルは、アプリケーションレイヤに置かれ、レングスインジケータは、関連するプロトコルデータユニット(PDU)の全長をバイトで示す。この数は、この例示的な実施形態において、4Kを超えない。PDUタイプインジケータは、それぞれがそれに関連づけられたそれ自体のフォーマットを有する256個のPDUタイプのうちのいずれか1つを示すために使用され得る。チェックサムバイトは、PDUのためのチェックサムを記憶する。

【0035】PDUは、256個のタイプのうちのいずれかであり、次のものを含む。

PDUタイプ: 00000000 ファンクション: パスキャッシング(Path Caching)

対応する可変長フィールドは以下のものを含む。

ソースノードNSAPアドレス: 20バイト

あて先ノードNSAPアドレス: 20バイト

1番目の中間ノードNSAPアドレス: 20バイト

...

k番目の中間ノードNSAPアドレス: 20バイト

...

最後の中間ノードNSAPアドレス: 20バイト

【0036】PDUタイプ: 00000001 ファンクション: パスキャッシングコンファメーション(Path Caching Confirmation)

対応する可変長フィールドは以下のものを含む。

20バイト

20バイト

1バイト

(値: YES=1/NO=0)

【0037】PDUタイプ: 00000002 ファン  
クション: パスリムーバル(Path Removal)  
対応する可変長フィールドは以下のものを含む。

ソースノードNSAPアドレス: 20バイト

ソースノードNSAPアドレス: 20バイト

あて先ノードNSAPアドレス: 20バイト

STS-1スロットの数: 1バイト

STS-1スロット番号: 先行するフィールドにより決定される長さ

【0039】PDUタイプ: 00000004 ファン 10  
クション: 帯域幅割当てコンファメーション

ソースノードNSAPアドレス: 20バイト

あて先ノードNSAPアドレス: 20バイト

レスポンス値: 1バイト

(YES=1/NO=0)

【0040】PDUタイプ: 00000005 ファン  
クション: 帯域幅割当て解除

ソースノードNSAPアドレス: 20バイト

あて先ノードNSAPアドレス: 20バイト

STS-1スロットの数: 1バイト

STS-1スロット番号: 先行するフィールドにより決定される長さ

【0041】PDUタイプ: 00000006 ファン  
クション: トンネルド(Tunneled)OSPF  
対応する可変長フィールドは以下のヘッダを含む。

ソースノードNSAPアドレス: 20バイト

あて先ノードNSAPアドレス: 20バイト

以下の5個のタイプメッセージを含むオリジナルOSPF  
メッセージで

ハロー (Hello)

データベースディスクリプション

リンクステートリクエスト

リンクステートアップデート

リンクステートAck

LDP-REQUEST: ラベルリクエストメッセージ

LDP-MAPPING: ラベルマッピングメッセージ

LDP-WITHDRAW: ラベルウィズドローメッセージ

LDP-RELEASE: ラベルリリースメッセージ

LDP-NAK: LDPノーティフィケーション

【0043】PDUタイプ: 00000008 ファン  
クション: トンネルドMPLSCRLDP  
対応する可変長フィールドは以下のヘッダメッセージを  
含む。

ソースノードNSAPアドレス: 20バイト

あて先ノードNSAPアドレス: 20バイト

以下のメッセージを含むオリジナルの制約ルーティング  
ラベルディストリビューションプロトコル(CRLDP)  
メッセージとともに、

CRLDP-REQUEST: ラベルリクエストメッセ  
ージ

CRLDP-MAPPING: ラベルマッピングメッセ 50

あて先ノードNSAPアドレス: 20バイト

【0038】PDUタイプ: 00000003 ファン  
クション: 帯域幅割当て

対応する可変長フィールドは以下のものを含む。

対応する可変長フィールドは以下のものを含む。

注: この仮定は、OSPF-2がサポートされているこ  
とである。OSPF-2は、例えば、J. Moy, "OSPF R  
2.0, IETF RFC2328, ftp://ftp.isi.edu/in-notes/rfc2  
328.txt April 1998, に記載されている。

【0042】PDUタイプ: 00000007 ファン  
クション: トンネルドMPLSLDP対応する可変長フ  
ィールドは以下のヘッダメッセージを含む。

ソースノードNSAPアドレス: 20バイト

30 あて先ノードNSAPアドレス: 20バイト

以下の5個のタイプメッセージを含むオリジナルLDP  
メッセージとともに、

40 【0044】この例示的な実施形態において、LDP中  
で使用されるラベルウィズドロー、ラベルリリースおよ  
びラベルノーティフィケーションメッセージは、CRL  
DPに対して直接的に使用され得る。

【0045】前述したように、本発明は、SONETま  
たはSDHを使用して採用され得る。説明の便宜および  
明瞭さのために、以下の例示的な実施形態は、STS-  
NのようなSONETに関連する用語を使用してSON  
ETについて説明されるが、当業者によりSDH実施形  
態に拡張され得る。

【0046】1. 初期化



1. 1 インタミディエートシステム—インタミディエートシステム (IS-IS) ベーストポロジオートディスカバリー (Intermediate System-Intermediate System (IS-IS) based topology Auto-discovery): 初期化において、各CAPがネットワークサービスアクセスポイント (NSAP) アドレスおよびIPアドレスでプロビジョンされたと仮定し、CAPノードはOSIスタックおよびTCP/IPスタックの両方をランさせる。これに対応して、各コアインタミディエートポイント (CIP) ノード (厳密にデジタルクロスコネクトシステム (DCS) でありSONET/SDHトランスポートの一部として働くSONET/SDHネットワークエレメント) が、NSAPアドレスでプロビジョンされ、OSIスタックのみをランさせる。

【0047】レベル2IS-ISがサポートされると、マルチプルエリアベースドインフラストラクチャが、サポートされ、フル20バイトNSAPが使用される。シングルリング、リングベースドメッシュインタコネクション、メッシュベースドトポロジなどのようないかなる対応が物理ネットワークトポロジも、新しいコンバージェンスプロトコルと共に使用され得る。本発明の側面において、この仮想自律的システムのトポロジは、IS-IS HELLOメッセージの交換に基づいてIS-IS隣接関係を解決することにより、そして、IS-ISリンクステートPDU (LSP) の交換に基づいてCAPおよびCIPを含む各ノードの物理トポロジを確立することにより、自動的に見いだされ得る。

【0048】1. 2 CAPパス確立

各CAPノードは、この仮想自律的システムに対するゲートウェイとして働く。各CAPは、インテリアゲートウェイプロトコル (IGP) またはエクステリアゲートウェイプロトコル (EGP) ルーティングメッセージのいずれかによりエクスターナルルーティングネイバー (external routing neighbors) との隣接関係を確立する。隣接関係の確立の後に、CAPは、ルーティングアップデートメッセージにより外部ルート到達可能性情報を得る。

【0049】各CAPペアは、内部論理パスを確立する。即ち、CAP間の1つまたは2つ以上の中間ノードがあり得るが、各CAPは、他のCAPのノードアドレスを見いだすことにより、他のCAPへの論理パスを確立する。論理パスを確立するために、IS-IS通信は、以下のように強化される。通常のIS-IS通信は、"Intermediate System To Intermediate System Intra-domain Routing Exchange Protocol For Use In Conjunction With The Protocol For Providing The Connectionless-mode Network Service (IS08473)", ISO DP 10589に記載されている。通常のIS-ISレベル2LSPは、図5に示されたフォーマットを示す。

【0050】本発明の原理によれば、図5のフォーマット

トが使用され、これにより現存するIS-ISスタックとのコンパチビリティを維持する。しかし、可変長フィールドは、各CAPがIS-IS LSPアップデートに基づいて他のCAPへそのIPアドレス情報を溢れ(flood)させることができるように拡張される。

【0051】IS-ISレベル2LSP可変長フィールドフォーマットは、図6に示されており、1つのバイトがコードを示し、1つのバイトがフィールドの長さを示し、残りが値を示す。

【0052】本発明の原理によれば、2つのCODE、CODE15およびCODE16が、図7の強化されたIS-ISレベル2LSP可変長フィールド図に示されているように、IP-V4およびIP-V6アドレスアナウンスメントをサポートするために使用される。CODE15は、IP-V4アドレスをアドバタイズするために使用され、CODE16はIP-V6アドレスをアドバタイズするために使用される。IS-IS LSPアップデートの期間の間に、各CAPは、強化されたIS-IS LSPに基づいてそのマネージメントIPアドレス (IP-V4またはIP-V6のいずれか) をアドバタイズし、各CIPは、強化されたIS-IS LSPフィールドを無視することになる。

【0053】強化されたIS-IS LSPの受信に回答して、CIPはPDUを通常通りに処理し、CAPは入ってくる情報に基づいてそのルーティングテーブルを追加的にリフレッシュすることになる。入ってくるデータが新しいIPアドレスを含む場合、CAPは、図8のバインディングテーブル図に示されているように、そのアドレスバインディングテーブルに新しいエントリを追加する。CAPにより受信される新しいアドレスが別のCAPのアドレスである場合、各CAPがそのIPマネージメントアドレスを別のCAPにアドバタイズするだけであるという過程で、プロセスが以下のように進む。

【0054】プロセス1、論理パス確立: (別のCAPノードBからIPアドレスをちょうど受信した現在のノードをAとする)

ステップ1: ノードBのIPアドレスに関連づけられた送信者のNSAPアドレスを見つける。

ステップ2: IS-ISルーティングテーブルルックアップに基づいてAとBとの間の物理パスを見だし、パス情報を記録する。

【0055】ステップ3: コンバージェンスプロトコルバスキャッシングメッセージをこの物理パスで送る。このメッセージ内で、ソースノードはAであり、あて先ノードはBであり、中間ノードのNSAPアドレスも含む。

ステップ4: パス上の各ノードに対して、そのようなバスキャッシングメッセージの受信の後、そのバス情報テーブルをリフレッシュする。AとBとの間にエントリがない場合、新しいエントリを追加し、そうでない場合、



次のノードのNSAP情報を対応するエントリに入れる。このバステーブルの各エントリは、図9のバステーブル図に示されたフォーマットでなければならない。

【0056】ステップ5：これが最終あて先であるかどうかを決定する。ノーである場合、それ自体のNSAPアドレスをバスキャッシングメッセージから取り除き、この修正されたメッセージを次のノードに送る。イエスである場合、ソースAからバス情報が受信されたことを記録する。それがAに到達するバスを有するかどうかをダブルチェックする。

【0057】イエスである場合、バスコンファメーションメッセージをソースAに送り返す。ノーである場合、受信されたメッセージ中に含まれるAのNSAPアドレスに基づいて、IS-ISルーティングテーブルルックアップに基づいてAに到達する物理バスを生成し、Aにそれ自体のバスキャッシングメッセージを送り（ソースノードはBであり、あて先ノードはAであり、それは中間ノードのNSAPアドレスを含む）、そしてバスコンファメーションメッセージをソースAに送り返す。

注：この場合において、各中間ノードは、BからAへのバス情報をまず受信し、その後受信される（フォワード）バスコンファメーションメッセージを処理することを保証される。

【0058】上記のプロセスにより、各CAPはそれ自体のアドレスバインディングテーブルを確立するだけでなく、他のCAPに到達するバス情報をも得る。

【0059】本発明の原理によれば、スタティックにプロビジョンされたインタネットチャンネルは、ダイナミック帯域幅マネージメントメカニズムによりバースト性のトラフィックを満足させるために使用される。ダイナミック帯域幅マネージメントメカニズムは、新しいヘッダテーブルおよびステータステーブルのようなリソースマネージメント構造および、CAPが様々な帯域幅リクエストに応答するために使用できるプロセスを含む。ダイナミック帯域幅マネージメントを提供するために使用されるメカニズムおよびプロセスは、図10および11との関連でより詳細に説明される。

【0060】CAPおよびCIPノードの両方に対して、1つの物理リンクが1つの帯域幅に関連づけられている場合、各リンクに関連づけられたリソーステーブルがある。WDMが使用される場合、各入に関連づけられたリソーステーブルエントリがある。このリソーステーブルの属性は、その隣接するもののアドレス情報、およびそのSTS-1支流（tributaries）の各々の物理的ステータスを含む。ここで、SONETネットワークエレメントに、STS-1レベルクロスコネクタキャパシティが備えられていることが仮定されている。

【0061】このリソーステーブルの取り扱い、新しいプロトコルを、現在のSONET/SDHのような単一波長オプティクスおよび多波長オプティクスを備えた

アプリケーションに対して適切なものにする。1番目の場合において、各リンクはリソースヘッダテーブルにおいて1つのエントリと関連づけられており、2番目の場合において、各波長は、リソースヘッダテーブルにおいて1つのエントリと関連づけられている。

【0062】各ノードに対するヘッダ情報は、図10に示されているようにテーブルとして組織され、このテーブル中の各エントリは、1つのリンクに関連づけられているかまたは1つの波長に関連づけられている。各エントリに対するフォーマットは、図10に示されている。このテーブル中の各エントリに対して、初期化時において、NSAPアドレスフィールドが、ネイバーのNSAPアドレスとしてプロビジョンされ、“利用可能な帯域幅”フィールドは、このリンク/波長の物理的容量またはそれがサポートできるSTS-1の数としてプロビジョンされる。そして、STS-1アレイポインタは、ランタイムにおいて初期化され、このヘッダテーブルエントリに関連づけられた対応するSTS-1アレイの物理的スターティングアドレスを指し示す。

【0063】本発明の原理によるコンバージェンスプロトコルベースドトランスポートネットワークは、現行のSONET装置とコンパチブルな装置を使用することができる。結果として、UPSRおよびBLSRのような現行のSONET復元アプローチが、直接的に適應され得る。また、各CAP中のルーティング/LSR機能で、最新のMPLSベースド復元アプローチが、SONETバス/ラインレベルの代わりにIPフローレベルにおける保護グラニュラリティ（granularity）をエンドユーザが好む場合に使用され得る。一般に、新しいトランスポートネットワークアーキテクチャは、様々な保護スイッチングソリューションを提供する。

【0064】STS-1ステータステーブルエントリのフォーマットは、図11に示されており、“フリーステータス”は、このSTS-1スロットが割り当てられたかどうかを示し、あて先NSAPアドレスはこの支流（tributary）の対応するあて先アドレスが何であるかを示し、“利用可能な帯域幅”は、このSTS-1スロット内の利用可能な帯域幅を示す。

【0065】本発明の原理によれば、ランタイムにおいて、各CAPは様々なフォーマットの帯域幅リクエストを受信することができ、これは、SNMPコマンドによりサービスプロバイダから明示的（explicitly）に得ることができ、またはMPLSラベルスイッチングパスセットアッププロセスから黙示的（implicitly）に得られる。これらのリクエストは、2つの異なるカテゴリ、即ち、帯域幅割当ておよび帯域幅割当て解除に分類され得る。各リクエストに関連づけられて、IPあて先アドレスおよび帯域幅を含む情報が提供される。

【0066】これに応じて、本発明の原理によるCAPは、以下に示すように帯域幅を割当てまたは割当て解除

10

20

30

40

50

することができる。説明されるプロセスのうちのいずれかが、明示的SNMPプロビジョニングコマンドとして働くために使用することができる。また、いずれのプロセスもダイナミック割当てをサポートするために使用され得る。ダイナミック割当ては、例えばMPLS明示的ルーティングのようなアプリケーションをサポートするために使用され得る。

【0067】STS-1レベル帯域幅ダイナミック割当て

注：ここで、帯域幅要求は、STS-1のN倍でなければならない。

プロセス2：CAPノードAを開始するため：

ステップ1：IPあて先アドレスを受信する。

ステップ2：アドレスバインディングテーブルにより対応するNSAPアドレスを見つける。そして、パステータブルに基づいて、次のCAPに到達するための物理パスを見つける。

【0068】ステップ3：パステータブルから得られた情報に基づいて次の中間ノードのNSAPアドレスを見つける。これをノードBとし、以下の条件を満足するリンクを見つけだすためにリソースヘッダテーブルをチェックする。このリンクのネイバーはB。このリンクの利用可能な帯域幅は必要とされる帯域幅を超える。そのようなリンクがない場合、ノーを応答する。

【0069】ステップ4：コンバージェンスプロトコル帯域幅割当てメッセージを次のノードに送る。メッセージはソースノードNSAPアドレス、あて先ノードNSAPアドレス、必要とされる帯域幅、および同定されたSTS-1スロットを含む。そして、次のノードからのコンファメーションメッセージを待つ。

ステップ5：次のノードから帯域幅割当てコンファメーションメッセージを得る。イエスである場合、帯域幅割当てを反映するように、対応するリソースヘッダテーブルおよび関連するSTS-1ステータステーブルを修正する。対応するフレーマ (framer) をプロビジョンする。イエスを応答する。ノーである場合、ノーを応答する。

【0070】中間CIPノードおよびあて先CAPノードに対して、

ステップ1：先行するノードからの帯域幅割当てメッセージの受信の後、それが最終あて先であるかどうかをチェックする。イエスである場合、先行するノードに帯域幅割当てコンファメーションメッセージによりイエスを応答し、STS-1のこのグループ内のトラフィックがパケットスイッチング機構により処理されるようにパスターミネーションとして対応するSTS-1をプロビジョンし、対応するフレーマをプロビジョンする。

【0071】ステップ2：そうでない場合、これは最終あて先ではない。パステータブルにより、次のNSAPアドレスを見いだし、これをNとする。

ステップ3：以下の条件を満足するリンクを見つけだすためにリソースヘッダテーブルをチェックする。このリンクのネイバーはNである。このリンクに対して利用可能な帯域幅は必要とされる帯域幅を超えない。そのようなリンクがない場合、帯域幅割当てコンファメーションメッセージによりノーを応答する。

【0072】ステップ4：コンバージェンスプロトコル帯域幅割当てメッセージを次のノードに送る。メッセージは、ソースノードNSAPアドレス、あて先ノードNSAPアドレス、必要とされる帯域幅、現在のノードが割り当てられたSTS-1番号を含む。そして、次のノードからのコンファメーションメッセージを待つ。

ステップ5：次のノードからコンファメーションメッセージを得る。イエスの場合、帯域幅割当てを反映するように対応するリソースヘッダテーブルおよび関連するSTS-1ステータステーブルを修正する。先行するノードにより特定されたSTS-1スロットとそれが割り当てたSTS-1スロットとの間に対応するクロスコネクタをセットアップし、この情報をクロスコネクタテーブルに記録し、イエスを応答する。ノーである場合、ノーを応答する。

【0073】プロセス3：STS-1レベル帯域幅割当て解除：

イニシエーティングCAPノードに対して：

ステップ1：IPあて先アドレスを受信する。

ステップ2：アドレスバインディングテーブルにより所定のIPアドレスの対応するNSAPアドレスを見つける。

ステップ3：パステータブルから得られた情報に基づいて、次の中間ノードにNSAPアドレスを見つける。リソースヘッダテーブルおよび関連するSTS-1ステータステーブルから対応するNSAPに到達するために割り当てられたSTS-1を見つけだし、所定の要求条件に基づいてその正確な数を選ぶ。

【0074】ステップ4：コンバージェンスプロトコル帯域幅割当て解除メッセージを次のノードに送る。メッセージは、ソースノードNSAPアドレス、あて先ノードNSAPアドレス、リリースされた帯域幅およびそれら同定されたSTS-1スロットを含む。

ステップ5：帯域幅割当て解除を反映するように対応するリソースヘッダテーブルおよび関連するSTS-1ステータステーブルを修正する。このグループのSTS-1を割当て解除するために、フレーマを再プロビジョンする。

【0075】中間CIPノードおよびあて先CAPノードに対して：

ステップ1：先行するノードからの帯域幅割当て解除メッセージの受信の後、それが最終あて先であるかどうかをチェックする。イエスである場合、入ってくるメッセージに含まれているSTS-1のフレーマを再プロビジ

ョンし、戻る。

ステップ2：そうでない場合、これは最終あて先でない。バステーブルにより、次のNSAPアドレスを見つけだし、それをNとする。

【0076】ステップ3：クロスコネクトテーブルに基づいて、対応するSTS-1クロスエントリを見つけ出す。対応する外へ出るSTS-1スロットエントリを見つけだし、これらのクロスコネクトプロビジョンを削除する。

ステップ4：コンバージェンスプロトコル帯域幅割当て解除メッセージを次のノードに送る。メッセージは、ソースノードNSAPアドレス、あて先ノードNSAPアドレス、必要とされる帯域幅およびこれらの同定されたSTS-1スロット番号を含む。そして、戻る。

【0077】以上に説明されたトポロジーディスカバリーおよびダイナミック帯域幅割当ては、最小のブロッキングを提供するエンドツーエンドマルチプロトコルラベル効果(MPLS)パスを確立するために使用される。QoSを強化されたMPLSをサポートするために、各CAPノードは、L. Anderson 等による"LDP Specification, available at <http://www.ietf.org/internet-drafts/draft-ietf-mpls-ldp-03.txt>, january 1999 に記載されたラベルディストリビューションプロトコル(LDP)およびB. Jamoussi 等による"Constraint-based LSP Setup Using LDP," <http://www.ietf.org/internet-drafts/draft-ietf-mpls-ldp-01.txt>, に記載されたコンストレインドリレーティングラベルディストリビューションプロトコル(CRLDP)をサポートすべきである。

【0078】この例示的な実施形態において、各CAPは、ラベルスイッチングルータ(LSR)として働く。CAPがコア(または中間)LSRとして配置されていると仮定する場合、エッジLSRの役割は、通常のIPヘッダをラベルに変換し、LDPパスセットアップメッセージを開始することである。

【0079】明示的ルーティングに基づくトラフィックエンジニアリングをサポートするために、CRLDPメッセージが、ラベルスイッチングパスを明示的にセットアップするために使用される。各CAPによりアドバタイズされる情報に基づいて、外部ルータが、この仮想自律システムについてのクリアなトポロジー情報を有することになる。結果として、外部ルータは、この仮想自律システムをトラバースするパスをセットすることができる。

【0080】ベストエフォートデリバリーを超えて、エンドツーエンドQoSラベルスイッチングをサポートするために、本発明の原理によるシステムは、明示的パスに沿って、対応するIPフローに関連づけられたトラフィックコントラクトを送る。結果として、パス上の各ノードは、事前に、トラフィックコントラクトリクエスト

をサポートできるかまたは否定するかを決定することができる。

【0081】エンドツーエンドQoSベースドLSPをセットアップするために使用されるプロセスが続く。CRLDPはハードステートインプリメンテーションに基づくので、RSVPの代わりにCRLDPが帯域幅を予約するために使用されると仮定されている。RSVPはソフトステートプロトコルインプリメンテーションに基づくので、周期的ステートリフレッシュが、非常に大きな帯域幅および演算を消費する可能性があり、したがってCRLDPが好ましい。

【0082】制御パスに関して、外部コアルータがIPフローのトラフィックリクエストを受信する場合、コアルータは、このリクエストを隣接するCAPに送り、明示的パスにおいて示された次のノードが隣のCAPである場合、これをAとする。そして、ノードAは、所定のトラフィックコントラクトおよびコネクションアドミッションコントロールメカニズムに基づいて、次のCAPに到達するために必要とされる等価な帯域幅を生成する。

【0083】そして、ノードAは、現存するポートがこのリクエストを満足するかどうかを決定するために、そのリソーステーブルを参照する。イエスの場合、ノードAは帯域幅を予約し、そのリクエストを次のノードに送る。一方、ノードAの現存するポートがリクエストを満足できない場合、ノードAはこのリクエスト満足させるために次のポートをセットアップするために必要とされるSTS-1の最も近いナンバーを計算し、以下に基づいて、内部パス(トリビュータリー)を決定する。

【0084】(1) 所定のリクエストに含まれる次のノード(CAP)情報

(2) アドレスバインディングテーブルおよびバステーブルに基づいてこのCAPに到達するための物理パス

(3) プロセス2に基づいて、このフローを受け入れるかこのリクエストを否定するかパスをプロビジョンする。このIPフローが移行されうる場合、新しいコンバージェンスプロトコルベースドトンネルドCRLDPメッセージが次のノードに送られる。

【0085】(4) 次のノードがあて先CAPではなくCIPである場合、コンバージェンスプロトコルベースドトンネルドCRLDPメッセージの受信の後、バステーブルに基づいて次のノード情報を得て、それが次のCAPに到達するまで、それを次のノードに送ることになる。

(5) 次のCAPは、リクエストをターミネートし、そのエグレス(egress)リンクを計算し、それがサポートできるかどうかを決定する。サポートできない場合、トンネルドCRLDPメッセージにより"ノー"をエントリCAPに送り返す。そうでない場合、該当する場合、リクエストを次のノード(ルータ/CAP)に送り、応答

を待つ。

【0086】(6) 次のCAPが外へ出ていくラベルを含む応答を得た後(ポジティブであることを仮定する、そうでない場合、ノーをエントリCAPに送り返す)、このIPフローに対するそれ自体のイングレスラベル(ingress label)を割当て、それが割り当てたラベルを含むトンネルDRLDPメッセージによりそれをエントリCAPに送る。

(7) あて先CAPにより割り当てられたエグレスラベルを含む応答をソースCAPが得た後、このIPフローに対するそれ自体のイングレスラベルを割当て、それを外部コアルータに送る。

【0087】上記の制御メッセージ交換により、本発明の原理によるシステムは、トランスポートコアをトラバースする入ってくるIPフローに対するLSPをセットアップする。LSPを確立した後、データフローは以下のようなになる。

【0088】(1) このIPフローのパケットは、コアルータからエントリCAP(これをAとする)に送られる。

(2) ラベルの正確なマッチングに基づいて、ノードAは、使用されるべき対応する内部チャネル/トリビュタリを見つけ、イングレスラベルをエグレスラベルにスワップし、次のノードがCIPである場合、ノードAはパケットを次のノードに送る。

【0089】(3) 中間SONETノードは、クロスコネク機構を介してパケットをこの内部チャネルを通して送る。この内部チャネルは、新しいコンバージェンスプロトコルシグナリングにより確立されたクロスコネク機構に基づくSONETパスである。

(4) あて先CAPは、ラベルの正確なマッチングにより内部チャネル/トリビュタリをターミネートし、イングレスラベルをエグレスラベルにスワップし、フローを次のLSRに送る。

【0090】このアプローチを使用することは、柔軟な帯域幅適用形バックボーンを提供する。また、最小のエンドツーエンド待ち時間が必要とされる。これは、エントリCAPおよびエクジットCAPのみが、レイヤ2ラベルスワッピングおよび関連するパケットフローキューイング(queueing)に参与するからである。また、中間CIPにおいて、レイヤ1パススルーは、最小の(決定的)トラバース待ち時間を与える。

【0091】本発明の原理によれば、帯域幅オンデマンドSONET/SDHトランスポートインフラストラクチャは、以下に示すように、新しいコンバージェンスプロトコルを使用して有効にされ得る。

【0092】1つの例示的な実施形態において、集中化されたリソースマネージメントアプローチを使用して、ネットワークマネージメントシステムは帯域幅割当てをトリガする。パスセットアップ期間の間に、これは以下

の情報をイニシエーティングノード(のみ)に送る。即ち、あて先アドレスおよび必要とされる帯域幅である。そして、コンバージェンスプロトコルは、パス上のノード間で帯域幅情報を交換するために使用される。上述したプロセス2および3は、このアプローチをサポートするために使用され得る。

【0093】ステップ1: NMSは、CAP S1のATM/IPサイドへ物理パスプロビジョニング情報を送る。これは、以下の情報を含む: (GUIインタフェースによると仮定する)

CAP S2のマネージメントATM/IPアドレス  
必要とされる帯域幅: この例において、OC-3Cを仮定する

両方の端部に対する論理リンクレイヤプロビジョニング情報(IPの場合において、フレームリレーまたはPPPプロビジョニング情報)

(両方の端部に対して) 論理ポートプロビジョニング情報

(両方の端部に対して) ATM/IPアドレス情報

ルーティング情報プロビジョニング(IPの場合に対して、OSPF/RIP/BGP、ATMの場合に対して、OPSF/PNNI)

【0094】ステップ2: CAP S1のATM/IPサイドが、CAP S2の所定のATM/IPマネージメントアドレス情報に基づいてそのアドレスバインディングテーブルを参照することにより、CAP S2の対応するマネージメントNSAPアドレスを見つける。CAP S2の対応するマネージメントNSAPアドレスを見つけることができない場合、否定的応答がNMSに送り返される。そうでない場合、ステップ3に進む。

【0095】ステップ3: CAP S1のATM/IPサイドは、プロビジョニング情報をそのSONETサイドに送る。これは、CAP S2の帯域幅要求条件情報およびマネージメントNSAPアドレスを含む。

【0096】ステップ4: CAP S1のSONETサイドは、まず、CAP内のインタコネク帯域幅がこの要求条件を満足するために十分であるかどうかを決定する。そうでない場合、それは、否定的応答をCAP S1のATM/IPサイドに送り返し、そして、この否定的応答はさらにNMSに送られる。

【0097】ステップ5: インタコネク帯域幅がこれに適合できる場合、確立されたSONETトポロジーに基づいて、CAP S1のSONETサイドは、CAP S2のSONETサイドに到達するためのパスを見つけだし、これはCAP S2に到達するために使用することができるそのエグレスポートをも見つけだす。もしできない場合、否定的応答をCAP S1のATM/IPサイドに送り返し、そしてこの否定的応答は、さらにNMSに送られる。

【0098】ステップ6: そして、CAP S1のSO

NETサイドは、同定されたエグレスポートのいずれかに対して、エグレスポート帯域幅がこの要求条件を満たすために十分であるかどうかを決定する。もし満足しない場合、それは、CAPS1のATM/IPサイドへ否定的応答を送り返し、そしてこの応答は、さらにNMSに送られる。

【0099】ステップ7：そのエグレスポート帯域幅がこの要求条件に適合できる場合、同定されたバス情報に基づいて、CAPS1のSONETサイドは、所定のSTS-1スロットを予約し、シグナリングメッセージフローを開始し、それを次のノードに送り、次のノードからの応答を待つ。このメッセージは、バス情報および予約されたSTS-1スロット情報を含む。(ソース/明示的ルーティングが使用されていると仮定している)

【0100】ステップ8：次のノードから受信された応答が肯定的である場合、CAPS1のSONETサイドは、対応するクロスコネクトをプロビジョンし、CAPS1のATM/IPサイドへ、適切なシグナリングにより物理的プロビジョニング情報を送る。そして、ATM/IPサイドは、物理レイヤ、論理リンクレイヤおよびATM/IPレイヤに対応するトリビュータリーをプロビジョンすることになり、そしてルーティングスタックを開始する。そして、肯定的応答は、NMSへ送り返される(この場合においてNavis)。次のノードから受信された応答が否定的である場合、CAPS1のSONETサイドは、STS-1スロット予約をキャンセルし、否定的応答をCAPS1のATM/IPサイドへ送り返し、そしてこの否定的応答は、NMSへさらに送られる。

【0101】分散型リソースマネージメントアプローチを使用して、各CAPノードは、アップツーデートトランスポートネットワークトポロジを有し、各CAPは、パフォーマンスモニタリング情報を集める。上記の情報に基づいて、これは、リンクの帯域幅が調節されるべきかどうかについてそれ自体で決定を行う。各CAPは、ルーティングテーブルルックアップに基づいてバスを確立し、パフォーマンスモニタリングに基づいて各確立されたトリビュータリーに対する帯域幅を再調節するための決定を行い、上述したコンバージェンスプロトコルプロセス2および3によりバス上の他のノードと帯域幅割当て/割当て解除情報を交換する。

【0102】図12において、システム1200は、ネットワークマネージメントシステム1202、それぞれATM/IPネットワーク1208および1210に接続された第1および第2のユーザ1204および1206を含む。ルータR1ないしR4は、ユーザ1204および1206を、ATM/IPネットワーク1208および1210を通してそれぞれCAP1212および1214に接続する。これは、SONET/SDHシステム1216のトランスポートフォスリティへのアクセス

を提供する。

【0103】そのような帯域幅割当てプロセスは、図13のシナリオを参照して以下に説明する。

ステップ1：NMSは、物理バスプロビジョニング情報をCAP1のATM/IPサイドに送る。これは、以下の情報を含む：(GUIインタフェースによると仮定している)

CAP2のマネージメントATM/IPアドレス  
必要とされる帯域幅：この場合において、OC-3Cを仮定している

両方の端部に対する論理リンクレイヤプロビジョニング情報 (IPの場合、フレームリレー、またはPPPプロビジョニング情報)

(両方の端部に対して) 論理ポートプロビジョニング情報

(両方の端部に対して) ATM/IPアドレス情報  
ルーティング情報プロビジョニング (IPの場合に対して、OSPF/RIP/BGP、ATMの場合に対して、OPSF/PNNI)

【0104】ステップ2：CAP1のATM/IPサイドは、CAP2の所定のATM/IPマネージメントアドレス情報に基づいてそのアドレスバインディングテーブルを参照することにより、CAPS2の対応するマネージメントNSAPアドレスを見つける。CAP2の対応するマネージメントNSAPアドレスが発見できない場合、否定的応答がNMSに送り返される。そうでない場合、ステップ3に進む。

【0105】ステップ3：CAP1のATM/IPサイドは、プロビジョニング情報をそのSONETサイドに送る。これは、CAP2の帯域幅要求条件情報およびマネージメントNSAPアドレスを含む。

【0106】ステップ4：CAP1のSONETサイドは、まず、CAP内のインタコネクト帯域幅がこの要求条件を満たすために十分であるかどうかを決定する。十分でない場合、これは、CAP1のATM/IPサイドへ否定的応答を送り返し、そして、この否定的応答はさらにNMSに送られる。

【0107】ステップ5：インタコネクト帯域幅がこれに適合できる場合、確立されたSONETトポロジに基づいて、CAP1のSONETサイドは、CAP2のSONETサイドに到達するためのバスを見つけだし、これは、CAP2に到達するために使用できるそのエグレスポートを見つけだす。もしできない場合、それは否定的応答をCAP1のATM/IPサイドに送り返し、そして、この否定的応答はNMSへさらに送られる。

【0108】ステップ6：そして、CAP1のSONETサイドは、いずれかの同定されたエグレスポートに対してエグレスポート帯域幅がこの要求条件を満たすために十分であるかどうかを決定する。十分でない場合、それは否定的応答をCAP1のATM/IPサイドに送

り返し、そしてこの応答はNMSへさらに送られる。

【0109】ステップ7：このエグレスポート帯域幅がこの要求条件に適合できる場合、同定されたバス情報に基づいて、CAP1のSONETサイドは、所定のSTS-1スロットを予約し、シグナリングメッセージフローを開始し、それを次のノードに送り、次のノードからの応答を待つ。このメッセージは、バス情報および予約されたSTS-1スロット情報を含む。(ソース/明示的ルーティングが使用されると仮定している)

【0110】ステップ8：次のノードから受信された応答が肯定的である場合、CAP1のSONETサイドは対応するクロスコネクをプロビジョンし、適切にシグナリングすることにより物理的プロビジョニング情報をCAP1のATM/IPサイドに送る。そして、ATM/IPサイドは、物理レイヤ、論理リンクレイヤおよびATM/IPレイヤを含む対応するトリビュータリーをプロビジョンすることになり、そして、ルーティングスタックをイニシエートする。そして、肯定的応答は、NMSへ送り返される(この場合においてNavis)。

【0111】次のノードから受信される応答が否定的である場合、CAP1のSONETサイドはSTS-1スロット予約をキャンセルし、そして否定的応答をCAP1のATM/IPサイドに送り返し、そしてこの否定的応答は、さらにステップ9ないし11においてNMSに送られる。

【0112】本発明の原理による別の例示的实施形態において、ネットワークは、プロトコルドリブプリソースマネージメントを使用する。このアプローチで、CRLDPまたはRSVPなどのような帯域幅予約プロトコルは、帯域幅割当てをトリガするために使用され得る。前述したように、これらのプロトコルは、(対応するQoSをサポートするために) 黙示的帯域幅要求条件および明示的バス情報を具現化する。コネクションアドミシジョンコントロールアルゴリズムに基づいて、エントリCAPノードは、黙示的帯域幅要求条件を等価な帯域幅に変換し、そして、入ってくるプロトコルメッセージに示されたバス上の他のノードと帯域幅割当て/割当て解除情報を交換する。

【0113】図14のフローチャートは、本発明の原理による初期化、CAPバスの確立、および帯域幅のダイナミックな割当てのプロセスを示す。本発明の原理による帯域幅の割当て解除もフローチャートに示されている。ステップ1400において、プロセスが開始し、ステップ1402に進み、コンバージェンスプロトコルを使用するノードが初期化される。初期化は、CAPにおけるOSIスタックおよびTCP/IPスタックの両方のランニングおよびCIPにおけるOSIスタックのランニングを含む。

【0114】プロセスは、ステップ1402からステップ1404に進み、ネットワークノードは、“IS-I

Sベーストポロジオーターディスカバリー”と呼ばれる上記のセクションに詳細に説明したオートディスカバリーのプロセスに従う。ステップ1404に続いて、プロセスはステップ1406に進み、各CAPペアは、それら自体の間に論理バスを確立する。この論理バスの確立は、ラベル交換バス可変長フィールドを使用する他のCAPに値するCAPのIPアドレス情報のフロディング(flooding)を必要とする可能性がある。論理バス確立プロセスは、“CPAバス確立”という名称の上記のセクションにより詳細に説明されている。

【0115】ステップ1406から、プロセスはステップ1408に進み、受信ノードは、受信ノードがCIPである場合、それを通常の方法で取り扱うことにより、本発明の原理に従って、IS-ISラベル交換バスメッセージの受信に応答する。さらに、CAPは、入ってくる情報に基づいて、そのルーティングテーブルをリフレッシュする。このプロセスは、図7との関連でより詳細に説明されている。ステップ1408から、プロセスはステップ1410に進み、CAPは、アドレスバインディングテーブルを確立し、他のCAPに到達するために必要なバス情報を得る。

【0116】このプロセスは、上述した“プロセス1”との関連でより詳細に説明されている。ステップ1410から、プロセスはステップ1412に進み、ノードは、それがあて先CAPに到達するまで、コンバージェンスプロトコル帯域幅割当てメッセージをノードからノードに渡す。ステップ1414において、バス上の帯域幅が、利用可能な場合、割当てられる。ステップ1412および1414のプロセスは、前述した“プロセス2”の説明においてより詳細に説明されている。

【0117】あるポイントにおいて、帯域幅割当て解除の必要性がある可能性があり、その場合において、プロセスはステップ1416へ進み、ノードはあて先CAPに到達するまで、ノードツーノードで、コンバージェンスプロトコル帯域幅割当て解除メッセージを渡す。ステップ1418において、適切に、帯域幅が割当て解除される。ステップ1416および1418のプロセスは、上述した“プロセス3”の説明においてより詳細に説明されている。全体のプロセスは、例えばメンテナンスまたはインストレーション動作の間に、例えばステップ1420において、エンドへ進み得る。

【0118】本発明の特定の実施形態の上記の記載は、例示のために示されたものである。これは、本発明を開示された正確な形式に限定することを意図したものではなく多くの修正および変形が上記の教示に照らして可能である。実施形態は、本発明の原理およびその実施を最適に説明するために選ばれかつ説明されたものであり、したがって、当業者は本発明を最適に利用可能である。本発明の範囲は、添付された特許請求の範囲によってのみ限定されることが意図されている。



## 【0119】

【発明の効果】以上説明したように、本発明によれば、SONET/SDHトランスポートネットワークの早い復旧、確かめられた安定性、低コストおよび低いトランスポート待ち時間を利用し、かつSONET/SDHトランスポートとIPトランスポートとの間のギャップを埋めて、動作コストを最小化しかつトラフィックエンジニアリングを容易にするシステムを提供することができる。

## 【図面の簡単な説明】

【図1】本発明の原理による通信ネットワークの概念的ブロック図

【図2】本発明の原理によるネットワーク要素の概念的ブロック図

【図3】本発明の原理による通信ネットワークのブロック図

【図4】新しいコンバージェンスデータプロトコルの原理によるプロトコルデータユニットを示す図

【図5】通常のIS-ISレベル2 LSPを示すブロック図

【図6】通常のIS-IS可変フィールド長フィールドを示すブロック図

【図7】本発明の原理による強化されたIS-ISレベル2 LSPを示す図

【図8】本発明の原理によるバインディングテーブル図

【図9】本発明の原理によるバステーブル図

【図10】本発明の原理によるヘッダエントリテーブル図

【図11】本発明の原理によるステータステーブル図

【図12】本発明の原理によるコンバージェンスプロトコルを使用する通信システムの概念的ブロック図

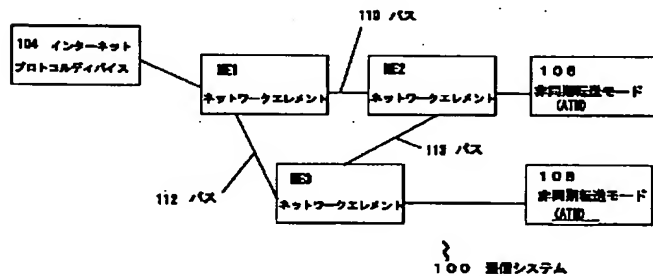
【図13】本発明の原理による帯域幅割当てプロセスを示すシナリオ図

【図14】本発明の原理による帯域幅割り当てと割り当て解除を含むプロセスを示すフローチャート図

## 【符号の説明】

- 100 通信システム
- 104 インタネットプロトコル (IP) デバイス
- 106 非同期転送モード (ATM) デバイス
- 108 ATMデバイス108
- 110、112、113 バス
- 200 ネットワーク要素
- 202 パケット交換コンポーネント
- 204 回線交換コンポーネント
- 206 インタネットプロトコルおよび非同期転送モードスイッチ
- 208、210 ラインカード
- 212 SONETリンク
- 214 回線交換入力/出力
- 215 回線交換シェルフ
- 216、218、220 機構
- 222 ローカルスイッチコア
- 224 シェルフコントロール
- 226 インタフェース
- 228 セントラルスイッチ機構
- 230 シェルフコントロール
- 234 セントラルスイッチコア

【図1】



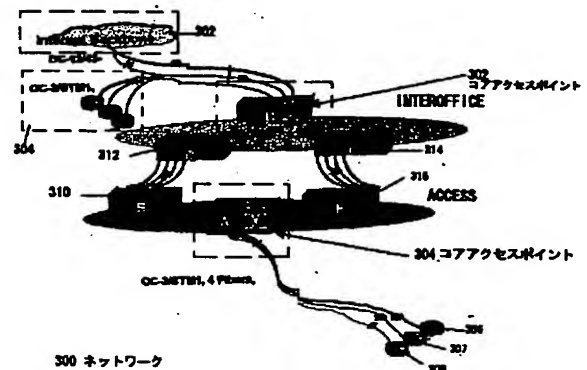
【図6】

バイト1	コード
バイト2	長さ
	値

【図9】

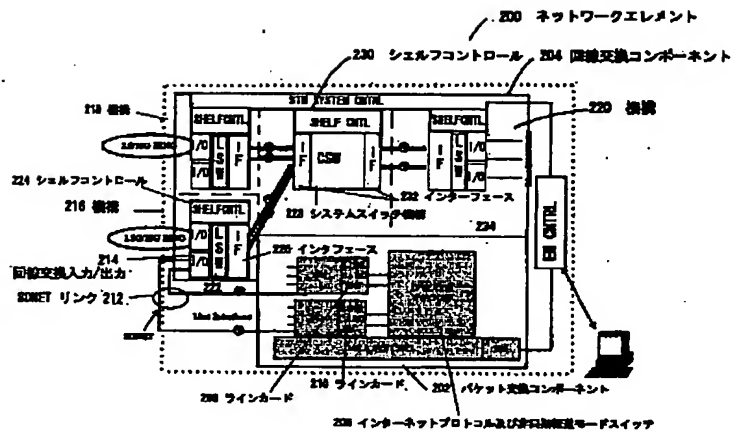
ソースノードのNSAPアドレス	あて先ノードのNSAPアドレス	次のノードのNSAPアドレス
-----------------	-----------------	----------------

【図3】





【図 2】



【図 4】

バイト1	コンバージェンスプロトコルディスクリミネータ	
バイト2	レンジサインジケータ	
バイト3	レンジサインジケータ(4ビット)	バージョン番号(4ビット)
バイト4	PDU タイプ	
バイト5	チェックサム	
	可変長フィールド	

【图 5】

バイト1	イントラドメインルーティングプロトコルディスクリミネータ
バイト2	レンジスインザキータ
バイト3	バージョン/プロトコル ID エクステンション
バイト4	ID長
バイト5	R   R   R   PDUタイプ
バイト6	バージョン
バイト7	予約済
バイト8	最大エリアアドレス
バイト9	PDU長
バイト10	PDU長
バイト11	残りライフタイム
バイト12	残りライフタイム
バイト13	LSP ID
バイト14	シーケンス番号
バイト15	シーケンス番号
バイト16	シーケンス番号
バイト17	シーケンス番号
バイト18	チェックサム
バイト19	チェックサム
バイト20	P   ATT   LSPDBOL   IS タイプ
	可変長フィールド

【图 8】

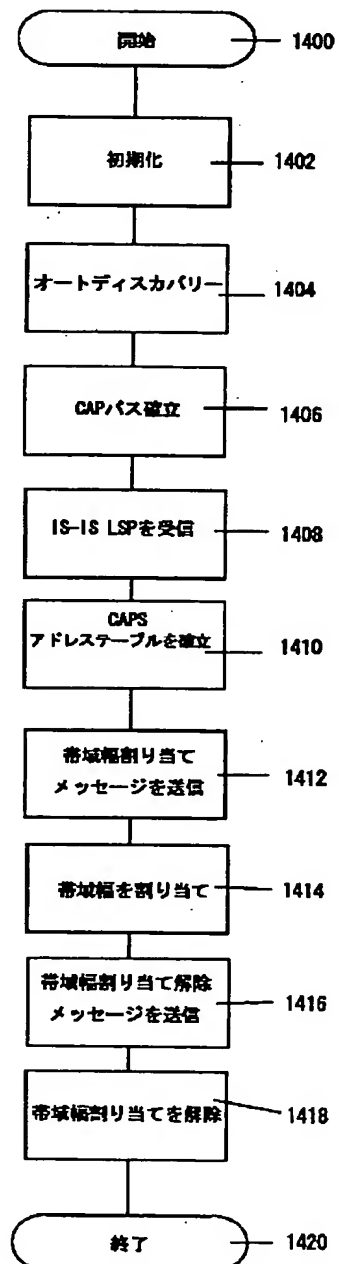
IP アドレス (IP-V4なら4バイト、IP-V6が サポートされていれば7バイト)	対応する NSAP アドレス (20バイト)
---	---------------------------

【図 7】

バイト1	コード = 15
バイト2	長さ = 4
	IP-V4 アドレス

バイト1	コード = 16
バイト2	長さ = 16
	IP-V6 アドレス

【图 14】



【図10】

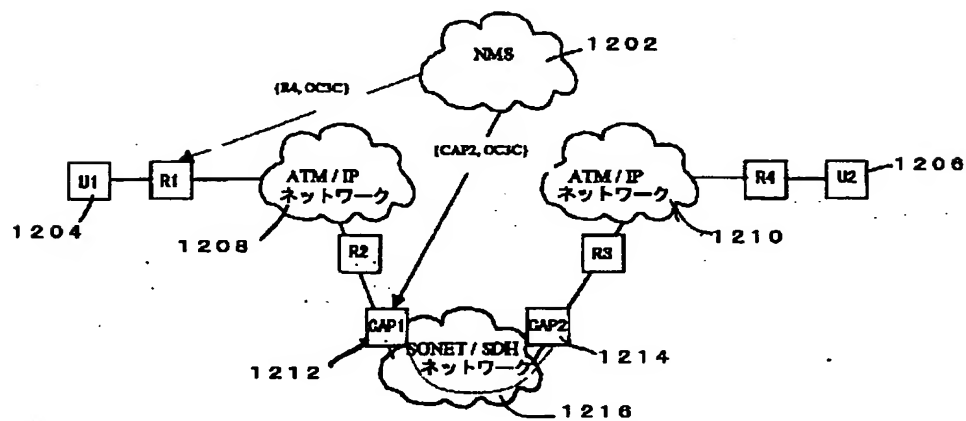
ネイバーの NSAP アドレス (20 バイト)	利用可能な帯域幅 (16バイト:ユニット:STS1)	STS-1 アレイポインタ (4バイト-8バイト、 32ビットモードか64ビ ットモードかに依存する)
--------------------------------	-------------------------------	--

【図11】

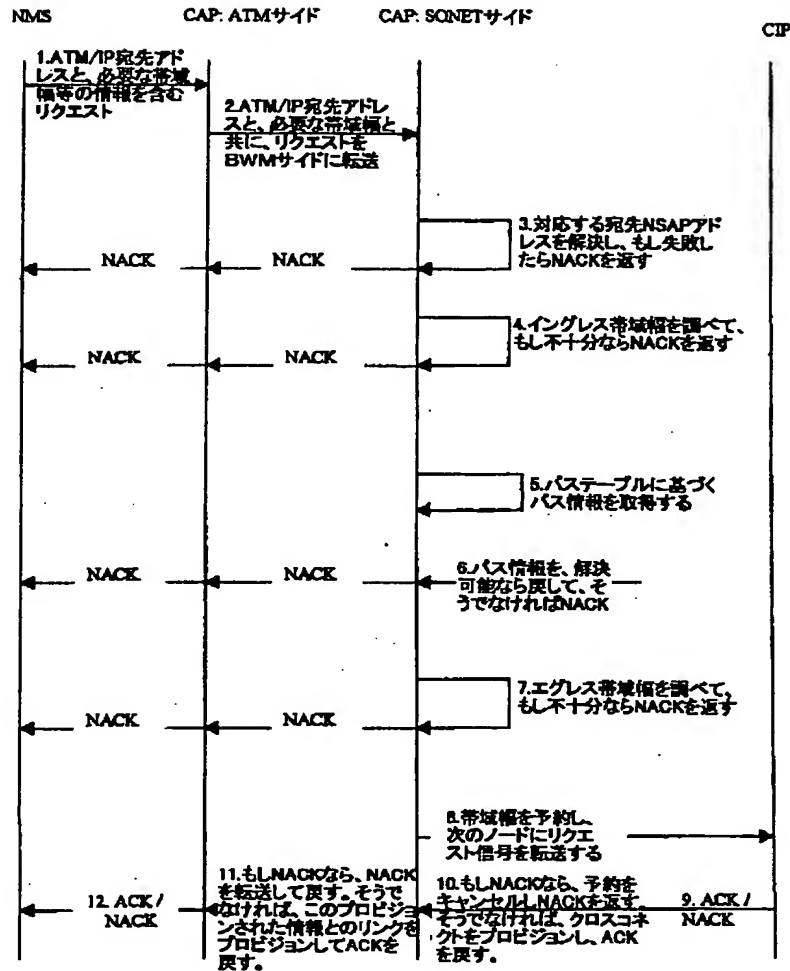
STS-1スロット番号	フリーステータス	あて先の NSAP アドレス	利用可能な帯域幅
-------------	----------	-------------------	----------

【図12】

1200



【図 13】



フロントページの続き

(71)出願人 596077259  
600 Mountain Avenue,  
Murray Hill, New Jersey 07974-0636 U. S. A.

(72)発明者 ヤン カオ  
アメリカ合衆国、01835 マサチューセツ、ブラッドフォード、フォレスト アクレス ドライブ 26B

## 【外国語明細書】

## 1. Title of Invention

Apparatus And Method For Integrated Telecommunications

## 2. Claims

Cao 3

32

## 1. A telecommunications system comprising:

a first network element which conforms to an OSI seven layer model; and

5 a second network element which conforms to an OSI seven layer model, each of the elements employing a convergence protocol whereby one of the layers of one of the network elements may communicate directly with a different layer of the other network element.

10 2. The system of claim 1 wherein the convergence protocol allows the transport layer of one network element to communicate directly with the service layer of another network element.

3. The system of claim 2 wherein the service layer is an internet protocol service layer.

4. The system of claim 2 wherein the transport layer is a synchronous optical network (SONET) transport layer.

15 5. The system of claim 2 further comprising a plurality of packet-switched devices and wherein said first and second network elements are connected to provide circuit-switched cross-connect for the packet switched devices.

6. The system of claim 5 wherein at least one of the packet-switched devices is an Internet protocol (IP) device.

20 7. The system of claim 5 wherein at least one of the packet-switched devices is an asynchronous transfer mode (ATM) device.

8. The system of claim 4 wherein a network element is responsive to a request from another network element using the convergence protocol for bandwidth along a link by granting bandwidth to the requesting network element.

Cao 3

33

9. The system of claim 8 wherein the request from one network element for bandwidth from another network element is effected via inter-layer, inter-element communications in a direct communications channel.
10. The system of claim 9 wherein a network element monitors layer three and layer one operations.
11. In a telecommunications system including packet-switched and circuit-switched network elements, a method comprising the steps of:
- (A) sending a message from one network element at one network layer; and
- (B) receiving the message of step (A) at a different network layer in a different network element.
12. The method of claim 11 wherein step (A) comprises the step of:
- (A1) communicating directly from the transport layer of one network element with the service layer of another network element.
13. The method of claim 12 wherein the service layer is an internet protocol service layer.
14. The method of claim 12 wherein the transport layer is a synchronous optical network (SONET) transport layer.
15. The method of claim 11 further comprising the step of:
- (C) providing circuit switched cross-connection for a plurality of packet-switched devices.
16. The method of claim 15 wherein step (C) comprises the step of:
- (C1) providing circuit switched cross-connection for an Internet protocol (IP) device.

Cao 3

34

17. The method of claim 15 wherein step (C) comprises the step of:

(C2) providing circuit switched cross-connection for an asynchronous transfer mode (ATM) device.

18. The method of claim 12 further comprising the steps of:

5 (D) sending a request from one network element to another for additional bandwidth; and

(E) granting additional bandwidth to a network element that requests additional bandwidth.

10 19. The method of claim 18 wherein step (D) is effected via inter-layer, inter-element communications in a direct communications channel.

20. The method of claim 19 further comprising the step of:

(F) a network element monitoring layer three and layer one operations; and

(G) the network element of step (F) adjusting the bandwidth of a logical pipe in response to changes in layer three and layer one operations.

15 21. A method of establishing communications paths in a hybrid, circuit-switched/packet-switched telecommunications system comprising the steps of:

(A) initializing nodes that provide an interface between SONET/SDH transport and external IP transport, by running both OSI stack and TCP/IP stack;

20 (B) initializing SONET/SDH network elements that are digital cross-connect, by running OSI stack; and

(C) auto-discovering the network topology using intermediate system-intermediate system (IS-IS) adjacency relationships.

Cao 3

35

22. The method of claim 21 wherein step (C) further comprises the step of:

(D) establishing the physical topology of each node within the telecommunications system based on the exchange of IS-IS Link State Protocol Data Units.

5 23. The method of claim 21 further comprising the step of:

(E) establishing a logical path between a pair of CAPs.

24. The method of claim 23 wherein step (E) further comprises the step of:

(E1) the refreshing of a routing table by a CAP in response to the reception of a IS-IS label switched path message.

10 25. The method of claim 24 wherein step (E) further comprises the step of:

(E2) a CAP's establishment of an address binding table and determining the path information required to reach another CAP.

26. The method of claim 23 further comprising the step of:

15 (F) passing a convergence protocol bandwidth allocation message to a destination CAP.

27. The method of claim 26 further comprising the step of:

(G) a destination CAP's allocation of bandwidth in response to the reception of a bandwidth allocation message as in step (F).

28. The method of claim 27 further comprising the step of:

20 (H) a destination CAP's de-allocation of bandwidth in response to the reception of convergence protocol de-allocation message.

### 3. Detailed Description of Invention



Cao 3

1

**FIELD OF THE INVENTION**

The invention relates to telecommunications systems and, more particularly, to the integration of circuit switched and packet switched telecommunications services.

**BACKGROUND OF THE INVENTION**

The explosive growth in the use of high-speed data transmission and the demand for differentiated data services strain the ability of existing telecommunications systems, a somewhat awkward marriage of circuit switched and packet switched systems, to meet the surging demand, both today and into the future. Conventional data communications systems which employ the publicly switched telephone network combine circuit switched TDM-based transport with best-effort packet switching, such as internet protocol (IP) switching, to effect data transmission in a manner that emphasizes maximum link utilization.

Telecommunications links may be established in a number of ways, with different advantages accruing to different methods of connection. The direct connection of two or more channels between two instruments, a connection that provides a user with exclusive use of the channels to exchange information, is referred to as a circuit switched, or line switched, connection. Circuit switching is a technique which yields highly reliable service and is particularly suitable for such "real time" communications applications as voice, in which the momentary loss of a channel is annoying, and repeated such losses are unacceptable. Circuit switching is also employed for highly reliable leased-line services. Electronic switching systems, such as the 5ESS may interconnect a multitude of telephone instruments through circuit switching, employing time division multiplexing (TDM), for example. In order to ensure that end-users receive the appropriate quality of service, the switches typically monitor and periodically test the activity of the trunk lines that carry the channels being switched. If a communications error occurs, the

Cao 3

2

switch may employ a "loopback" to isolate, or determine the exact location of, the system component that caused the error. Once the failure is isolated, the system may reconfigure itself so that data may be routed around the failed system component, through a loopback, or take other corrective measures. TDM transport networks provide an assured level of performance and reliability. Technologies, such as Synchronous Optical Network/Synchronous Digital Hierarchy (SDH/SONET) may be employed in the transport infrastructure to provide high-capacity transport, scalable to gigabit per second rates, with excellent jitter, wander, and error performance for voice connections and leased-line applications. SONET/SDH self-healing rings enable service-level recovery within tens of milliseconds following network failures.

Packet switching may be employed to maximize the utilization of telecommunications links, such as leased circuit-switched lines. With the packet switching approach, data is transmitted in packets, and the communications channel is only occupied for the duration of a packet's transmission. After the transmission, the channel is available for use by packets being transferred between other instruments. The links are statistically multiplexed to achieve maximum link utilization and are typically carried on leased circuits through the TDM transport network. Packet-switched systems often employ Internet Protocol (IP) transport methods to route packets from a source to a destination. Such systems generally employ "best-effort" techniques to deliver the packets and they generally lack the means to guarantee high reliability and predictable performance. Although statistical multiplexing yields high link utilization, the best-effort service provided by IP data networks is accompanied by unpredictable delay, jitter, and packet loss.

Although existing IP data networks provide excellent connectivity, they do not enable controllable distribution of network resources among the various service providers employing the IP data network to provide data transmission services to end users. That is, through a provisioning process that typically requires the intervention of craft workers, a process that takes place relatively

Cao 3

3

infrequently, the TDM-based transport services provide fixed bandwidth communications channels for each service provider. Since packet data traffic is inherently irregular, with bursts of heavy utilization followed by periods of relative inactivity, the fixed bandwidth "pipes" of TDM transport limit the flexibility with which transport users, such as IP data transport service providers, can respond to variation in demand from their end-users. Service providers must rely upon the good behavior of end users, who must scale back their transmissions during periods of heavy congestion. The well-known "tragedy of the commons" teaches that such reliance on the cooperative good behavior of a large group of users may not be a reasonable approach in the long run. This is made particularly clear in light of some applications, such as streaming video, that tend not to cooperate in scaling back traffic during periods of high usage.

Additionally, since the service layer and transport layer are separated in such a multi-layered dual architecture (circuit switched transport/IP) telecommunications system, transport management is segregated into separate operation and maintenance functions. In turn, this segregation of operations and maintenance functions generally requires the coordination of separate organizations and renders the process of end-to-end provisioning of a channel a difficult task. Such provisioning requires an inordinate amount of expertise and time, complicates the task of traffic engineering, and, as a result, the quality of service suffers while, at the same time, the transport core is underutilized.

A system and method which provide network infrastructure support to provide differentiated service guarantees and corresponding service level agreements to service providers, while taking advantage of both the high quality of service provided by circuit switched systems and the high utilization afforded by packet switched systems, would be highly desirable. Additionally, it would be highly desirable for such a system to dramatically increase, and maximally share, backbone network infrastructure capacity, and provide sophisticated service differentiation for emerging data applications, at least in part by dynamically managing the transport core bandwidth.

Cao 3

4

In short, a system which exploits the fast restoration, proven stability, low cost and low transport latency of SONET/SDH transport networks and bridges the gap between SONET/SDH transport and IP transport to thereby minimize operation cost and facilitate traffic engineering, would be highly desirable.

## 5 SUMMARY

A telecommunications system in accordance with the principles of the present invention includes a convergence protocol that provides efficient inter-layer communications between nodes within the network. For example, the convergence protocol permits the transport layer of one network node to  
10 communicate directly with the service layer of another network node. In an illustrative embodiment, the service layer is configured as an internet protocol (IP) layer, which may employ asynchronous transfer mode switching, and the transport layer is a SONET transport layer. Each of the nodes may be a telecommunications element such as an add-drop-multiplexers (ADM) or a digital  
15 cross connect, for example. This inter-layer communication is achieved through a layer-1 pass through operation.

In one aspect of the invention, SONET bandwidth provisioning, which heretofore has been static, is rendered flexible. Additionally, the division between transport and service is eliminated. This increased flexibility, in turn, provides  
20 support for differentiated service guarantees and corresponding service level agreements. The system provides the high quality of service typical of circuit switched systems, and, at the same time, features the high utilization of a packet switched system. A system in accordance with the principles of the present invention merges IP transport and SONET/SDH transport through a novel  
25 convergence protocol. The flexibility provided by this approach permits dynamic management of a link's internal transport bandwidth thereby accommodating bursty traffic at minimal cost, with little end-to-end latency. For example, given one channelized interface, such as a 4OC12C based OC48 link, with conventional static provisioning the termination of each tributary is fixed for extended periods,

Cao 3

5

often months. But, in accordance with the principles of the present invention, the termination point of a tributary may be dynamically altered to accommodate a service provider's run-time needs. Consequently, the potential re-configuration period may be reduced from months to minutes, or even seconds.

5 In accordance with the principles of the present invention a communications network employs a protocol hereinafter referred to as the convergence protocol, which encapsulates the SONET/SDH transport details and which permits internal bandwidth to be managed on the fly to accommodate varying service demands. As a result, in general, data service providers need no  
10 concern themselves with the internal details of transport network infrastructure because of the smooth interface between external Internet backbone and SONET/SDH transport backbone. The SONET/SDH transport backbone may adaptively rearrange its internal bandwidth thus permitting automated end-to-end provisioning. One or more network elements are configured as core access points  
15 (CAPs) to provide an interface between SONET/SDH transport and external IP transport.

#### BRIEF DESCRIPTION OF THE DRAWINGS

The above and further features, aspects, and advantages of the invention will be apparent to those skilled in the art from the following detailed description,  
20 taken together with the accompanying drawings in which:

Figure 1 is a conceptual block diagram of a communications network in accordance with the principles of the present invention;

Figure 2 is a conceptual block diagram of a network element in accordance with the principles of the present invention;

25 Figure 3 is block diagram of a communications network in accordance with the principles of the present invention;

Cao 3

6

Figure 4 is a diagram of a protocol data unit in accordance with the principles of the new convergence data protocol;

Figure 5 is a block diagram illustration of a conventional IS-IS level 2 LSP;

5 Figure 6 is a block diagram illustration of a conventional IS-IS variable field length field;

Figure 7 is a diagram of an enhanced IS-IS level 2 LSP in accordance with the principles of the present invention;

Figure 8 is a binding table diagram in accordance with the principles of the present invention;

10 Figure 9 is a path table diagram in accordance with the principles of the present invention;

Figure 10 is a header entry table diagram in accordance with the principles of the present invention;

15 Figure 11 is a status table diagram in accordance with the principles of the present invention;

Figure 12 is a conceptual block diagram of a communications system which employs a convergence protocol in accordance with the principles of the present invention; and

20 Figure 13 is a scenario diagram of a bandwidth allocation process in accordance with the principles of the present invention.

#### DETAILED DESCRIPTION

In the conceptual block diagram of Figure 1 a telecommunications system 100 including both packet switched and circuit switched components operates in accordance with the principles of the present invention. At the transport layer one

Cao 3

7

or more SONET/SDH network elements (NE1, NE2, and NE3 in this illustrative embodiment) provide the circuit-switched cross connect for packet switched devices such as internet protocol (IP) device 104, asynchronous transfer mode (ATM) device 106, and and ATM device 108. The IP device 104 and ATM devices 5 106 and 108 may be a combination of one or more devices that could be arranged in a ring topology for example. The system 100 is shown with only a limited number of devices in order to simplify the exposition and understanding of the invention. Two cross-connect paths, 110 and 112, between the IP device 104 and ATM device 106 and between the IP device 104 and the ATM device 108, 10 respectively, provide the specific circuit switched paths across the SONET/SDH transport layer. Each path may transit a plurality of SONET/SDH network elements and may share the bandwidth through each of those elements with other paths not shown in Figure1. A service provider may provision a portion of the bandwidth on each of the paths 110 and 112 for use by customers who wish to 15 transmit data between the IP device 104 and the ATM device 106 and between the IP device 104 and the ATM device 108, for example.

The layers discussed hereinafter refer to OSI layers, which are known and discussed, for example, by Ming-Chwan Chow, "Understanding SONET/SDH" ANDAN Publisher, Holmdel, NJ pages 2-31 through 2-32, which is hereby 20 incorporated by reference. Unlike conventional systems in which inter-layer communications are restricted to intra-element communications, a telecommunications system in accordance with the present invention may perform inter-layer communications between network elements. That is, the service layer of a network element such as NE1 may communicate with the transport layer of network element NE2. As will be described in greater detail below, this 25 communications path permits a network in accordance with the principles of the present invention to re-provision the paths 110 and 112 to accommodate changing bandwidth demand along the paths. This is in contrast with a conventional system wherein the provisioning is static and manual. That is, in a conventional system, 30 a service provider, such as AOL, would have to request re-provisioning by the



Cao 3

8

transport service provider, such as AT&T, and which would require the intervention of craft workers. Such re-provisioning and would typically take place only a few times each year.

5 In accordance with the principles of the invention if, for example, the path 110 is provisioned for a given transmission rate and a surge in demand occurs for the link between NE1 and NE2, the network element NE1 may directly request of the network element NE3 sufficient bandwidth to supply data to NE2 along paths 112 and 113 from NE1 to NE2. This immediate provisioning is accomplished through a direct inter-layer and inter-element communications in the direct  
10 communications channel (DCC) as described in greater detail below.

Each of the network elements NE1, NE2, and NE3, of Figure 1 could operate as a CAP, thereby permitting SONET transport for packet-switched networks such as those which employ Internet Protocol. A front-end aggregating router may be consolidated with an entry CAP, so packet level traffic may be  
15 served by an entry CAP. Additionally, because each CAP has the capability to monitor both layer-3 and layer-1 operations, each CAP also has the ability to re-arrange the size of individual logical-pipes "on the fly" (through exchanging Convergence Protocol message with other nodes). Consequently, the inventive system can adaptively adjust the bandwidth to optimize pipe-utilization, and, at the  
20 same time, improve service for best effort delivery traffic based on each CAP's performance monitoring. It may also adjust the bandwidth arrangement based on input from a centralized network management system which collects the global network traffic statistics.

The conceptual block diagram of Figure 2 illustrates the interconnections  
25 between the packet-switched 202 and circuit switched 204 components of a network element 200 in accordance with the principles of the present invention. The packet-switched component 202 includes an internet protocol and asynchronous transfer mode switch 206 that is operatively connected to line cards 208 and 210. The line cards 208 and 210 are connected through SONET links 212

Cao 3

9

to the circuit switched section 204 through the circuit switched Input/output 214 of a circuit switched shelf 216. The circuit switched section 202 may include a plurality of local switch fabrics, or circuit switched shelves, such as fabrics 216, 218 and 220. Each of the fabrics may include I/O, such as I/O 214 and 216 of the fabric 216, a local switch core 222, shelf control 224, and an interface 226. The interface 226 provides communications between the local switch fabric and a central switch fabric 228. The central switch fabric 228 includes shelf control 230, an interface 232 for each of the local fabrics, and a central switch core 234.

A network element such as the network element 200 of Figure 2 may be employed as an interface between circuit switched and packet switched networks. When the network element is used in this manner, it will be referred to as a core access point, hereinafter. The network 300 illustrated in the conceptual block diagram of Figure 3 includes core access points 302 and 304 which provide an interface between the IP transmissions of elements 306 and 308 and the SONET transport of elements 310, 312, 314, and 316. That is, transmissions among the network elements 310, 312, 314, 316, 302, and 304 are SONET/SDH transport transmissions, and transmissions among network elements 302, 304, 306, 307, and 308 are IP transmissions.

In accordance with the principles of the present invention, network elements that operate as core access points do so in conjunction with a new protocol that will be referred to hereinafter as the "convergence protocol". The convergence protocol encapsulates the OSI stack. The convergence protocol data unit format is illustrated in Figure 4. The data unit comprises a five byte header and variable length fields. The header includes a one-byte convergence protocol discriminator, a one-byte length indicator, one byte that is split between length indication and version number, a one-byte protocol data unit type indicator, and a checksum. Each protocol stack has its own layer structure. For example, in OSI there are seven layers: Physical, Link, Network, Transport, Session, Presentation, and Application. In this illustrative embodiment, the new protocol is positioned at the Application layer and the length indicator indicates the total

Cao 3

10

length of the associated protocol data unit (PDU), in bytes. This number may not exceed 4K in this illustrative embodiment. The PDU type indicator may be used to denote any one of 256 PDU types, each of which has its own format associated with it. The checksum byte is stores the checksum for the PDU.

- 5 As noted, the PDU may be any of 256 types, including:

PDU Type: 00000000

Function: Path Caching

The corresponding Variable Length Fields include the following:

Source node NSAP address: 20 bytes

Destination node NSAP address: 20 bytes

10 The 1<sup>st</sup> intermediate node NSAP address: 20 bytes

...

The k<sup>th</sup> intermediate node NSAP address: 20 bytes

...

The last intermediate node NSAP address: 20 bytes

15

PDU Type: 00000001

Function: Path Caching Confirmation

The corresponding Variable Length Fields include the following:

Source node NSAP address: 20 bytes

Destination node NSAP address: 20 bytes

20 Response value: 1 byte

(Value: YES = 1 / NO = 0)

Cao 3

11

PDU Type: 00000002

Function: Path Removal

The corresponding Variable Length Fields include the following:

Source node NSAP address: 20 bytes

Destination node NSAP address: 20 bytes

5

PDU Type: 00000003

Function: Bandwidth Allocation

The corresponding Variable Length Fields include the following:

Source node NSAP address: 20 bytes

Destination node NSAP address: 20 bytes

10

No. of STS-1 Slots: 1 byte

STS-1 Slot No.: length determined by previous field

PDU Type: 00000004

Function: Bandwidth Allocation

Confirmation

15 The corresponding Variable Length Fields include the following:

Source node NSAP address: 20 bytes

Destination node NSAP address: 20 bytes

Response value: 1 byte

(YES = 1 / NO = 0)

Cao 3

12

PDU Type: 00000005

Function: Bandwidth De-allocation

The corresponding Variable Length Fields include the following:

	Source node NSAP address:	20 bytes
	Destination node NSAP address:	20 bytes
5	No. of STS-1 Slots:	1 byte
	STS-1 Slot No.:	length determined by previous

field

PDU Type: 00000006

Function: Tunned OSPF

10 The corresponding Variable Length Fields includes the following header:

Source node NSAP address:	20 bytes
Destination node NSAP address:	20 bytes

With original OSPF message which includes the following five type messages:

	Hello
15	Database Description
	Link State Request
	Link State Update
	Link State Ack

20 Note: The assumption is that OSPF-2 is supported. OSPF-2 is discussed, for example, in J.Moy, "OSPF R2.0, IETF RFC2328, <ftp://ftp.isi.edu/in-notes/rfc2328.txt> April 1998, which is hereby incorporated by reference in its entirety.

Cao 3

13

PDU Type: 00000007

Function: Tunnelled MPLS LDP

The corresponding Variable Length Fields includes the following header message:

Source node NSAP address: 20 bytes

5 Destination node NSAP address: 20 bytes

with the original LDP message which includes the following five type messages:

10 LDP-REQUEST: Label Request Message  
 LDP-MAPPING: Label Mapping Message  
 LDP-WITHDRAW: Label Withdraw Message  
 LDP-RELEASE: Label Release Message  
 LDP-NAK: LDP Notification

PDU Type: 00000008

Function: Tunnelled MPLS CRLDP

15 The corresponding Variable Length Fields includes the following header message:

Source node NSAP address: 20 bytes

Destination node NSAP address: 20 bytes

20 with the original constrained based routing label distribution protocol (CRLDP) message which includes the following messages:

CRLDP-REQUEST: Label Request Message  
 CRLDP-MAPPING: Label Mapping Message

In this illustrative embodiment, the Label Withdraw, Label Release and Label Notification messages used in LDP may be used for CRLDP directly.

Cao 3

14

As previously mentioned the invention may be employed using SONET or SDH. For the convenience and clarity of exposition, the following exemplary embodiments will be described in terms of SONET, using SONET-related terminology, such as STS-N, but may be extended to SDH embodiments by one of ordinary skill in the art.

#### 1. Initialization

##### 1.1 Intermediate System-Intermediate System (IS-IS) based topology Auto-discovery:

At initialization, assuming each CAP node has been provisioned with network service access point (NSAP) address and IP address, the CAP node runs both OSI stack and TCP/IP stack. Correspondingly, each core intermediate point (CIP) node (a SONET/SDH network element that is strictly digital cross-connect system (DCS) and serves as part of SONET/SDH transport) is provisioned with NSAP address, and runs only OSI stack. With Level 2 IS-IS supported, multiple-area based infrastructure is supported and the full 20-byte NSAP is used. Any variety of physical network topology, such as single-ring, ring-based mesh interconnection, mesh based topology, etc., may be employed with the new, convergence, protocol. In one aspect of the invention the topology of this virtual autonomous system may be auto-discovered by establishing the IS-IS adjacency relationship based on an exchange of IS-IS HELLO message, then, establishing the physical topology of each node, including CAP and CIP, based on the exchange of IS-IS Link State PDU (LSP).

##### 1.2 CAP path establishment

Each CAP node serves as gateway for this virtual autonomous system. Each CAP establishes neighbor relationships with external routing neighbors through either an interior gateway protocol (IGP) or exterior gateway protocol (EGP) routing message. After establishing neighbor relationships, the CAP then obtains external route reachability information through a routing update message.



Cao 3

15

Each CAP pair establishes an internal logical path. That is, although there may be one or more intermediate nodes between CAPs each CAP establishes the logical path to the other CAP by discovering the other CAP's node address. To establish the logical path, IS-IS communication is enhanced as follows. Conventional IS-IS communication is described in, "Intermediate System To Intermediate System Intra-domain Routing Exchange Protocol For Use In Conjunction With The Protocol For Providing The Connectionless-mode Network Service (ISO8473)", ISO DP 10589, which is hereby incorporated by reference in its entirety. Conventional IS-IS Level 2 LSP exhibits the format illustrated in Figure 5.

10 In accordance with the principles of the present invention, the format of Figure 5 is employed, thereby maintaining compatibility with the existing IS-IS stack. However, the Variable Length Fields are expanded so that each CAP can flood its IP address information to the other CAPs based on IS-IS LSP update.

15 The IS-IS level 2 LSP Variable Length Field format is illustrated in the diagram of Figure 6, in which one byte describes the code, one byte describes the length of the field and the remainder, the value.

In accordance with the principles of the present invention, two CODEs, CODE 15 and CODE 16, are used to support IP-V4 and IP-V6 address announcement, as illustrated in the enhanced IS-IS level 2 LSP variable length field diagram of Figure 7. CODE 15 is used to advertise IP-V4 address and CODE 16 is used to advertise IP-V6 address. During the period of IS-IS LSP update, each CAP will advertise its management IP address (either IP-V4 or IP-V6) based on enhanced IS-IS LSP, and each CIP will ignore the enhanced IS-IS LSP field. In response to the reception of an enhanced IS-IS LSP a CIP will process the PDU conventionally, a CAP will additionally refresh its routing table, based on the incoming information. If the incoming data includes a new IP address, the CAP adds a new entry to its address binding table, as indicated in the binding table diagram of Figure 8. If the new addressed received by a CAP is the address of

Cao 3

16

another CAP, the process, in which the assumption is that each CAP will only advertise its IP management address to another CAP, proceeds as follows:

**PROCESS 1, Logical Path Establishment:**

(Denoting the current node as A, which just received IP address from another CAP node: B)

5

Step 1: Find the sender's NSAP address associated with IP address of node B

Step 2: Discover the physical path between A and B based on IS-IS routing table look-up, record the path information

10

Step 3: Send the Convergence Protocol Path Caching message down this physical path, inside this message, source node is A, destination node is B, it also contains the NSAP address of intermediate nodes.

15

Step 4: For each node along the path, after receiving such Path Caching message, refresh its path information table: if there is no entry between A and B so far, add new entry; otherwise fill next node's NSAP information in the corresponding entry. Each entry of this path table should be of the format illustrated in the path table diagram of Figure 9.

Step 5: Determine whether it's the final destination or not:

If NO, strip its own NSAP address from Path Caching message, and pass this modified message to next node.

20

If YES, record that it has received path information from Source A. Double check whether it has a path to reach A:

If YES, send the Path Confirmation message back to Source A.

25

If No, based on A's NSAP address as contained in the received message, generate physical path to reach A based on IS-IS routing table look-up, send its own Path Caching message to A (source node is B, destination node is A, it

Cao 3

17

also contains the NSAP address of intermediate nodes), then send Path Confirmation message back to Source A.

Note: In this case, each intermediate node is guaranteed to receive path information from B to A first, then process (forward) Path Confirmation message received later.

5

Through the above process, each CAP not only establishes its own address binding table, but also obtains the path information to reach the other CAPs.

10

In accordance with the principles of the present invention, statically-provisioned internet channels are used to satisfy bursty-traffic through a dynamic bandwidth management mechanism. The dynamic bandwidth management mechanism includes resource management structures, such as a new header table and status table, and processes that a CAP may use to respond to various bandwidth requests. The mechanisms and processes used to provide dynamic bandwidth management will be discussed in greater detail in relation to Figures 10 and 11.

15

For both CAP and CIP nodes, there is resource table associated with each link in case one physical link is associated with one wavelength. In case WDM is deployed, there is resource table entry associated with each  $\lambda$ . The attribute of this resource table includes the address information of its neighbor, and the physical status of each of its STS-1 tributaries, here assuming the SONET network element is equipped with STS-1 level cross-connect capacity.

20

This treatment of resource tables makes the new protocol suitable for application with single wavelength based optics (such as nowadays' SONET/SDH) and with multiple wavelength based optics. In the 1st case, each link is associated with one entry at resource Header Table, and in the 2nd case, each wavelength is associated with one entry at resource Header Table.

25

Cao 3

18

The header information for each node is organized as a table, as illustrated in Figure 10, and each entry in this table is associated with one link or associated with one wavelength. The format for each entry is illustrated in Figure 10. For each entry in this table, at initialization time, NSAP address field is provisioned as the neighbor's NSAP address, the Available Bandwidth field is provisioned as the physical capacity of this link/wavelength, or the number of STS-1 it can support. And STS-1 Array Pointer is initialized at run-time which points to the physical starting address of corresponding STS-1 array associated with this header table entry.

10 A Convergence Protocol based transport network In accordance with the principles of the present invention may employ equipment is compatible with and existing SONET equipment. Consequently, the existing SONET restoration approach such as UPSR and BLSR can be directly applied. Additionally, with a routing/LSR function in each CAP, the latest MPLS based restoration approach  
15 may also be used if an end-user prefers protection granularity at IP flow level instead of SONET path/line level. So in general, the new transport network architecture offers a variety of protection switching solutions.

The format of STS-1 Status Table entry is illustrated in the diagram of Figure 11, in which "Free Status" indicates whether this STS-1 slot has been  
20 allocated, the destination NSAP Address indicates what's the corresponding destination address for this tributary, and Available Bandwidth indicates the available bandwidth inside this STS-1 slot.

In accordance with the principles of the present invention, at run-time each CAP may receive various formats of Bandwidth Request, which may come  
25 explicitly from service provider through SNMP command, or implicitly from MPLS label switching path setup process. These requests may be classified into two different categories: bandwidth allocation and bandwidth de-allocation. Associated with each request, information including IP Destination Address and bandwidth is provided. In response, a CAP in accordance with the principles of the present

Cao 3

19

invention may allocate or de-allocate bandwidth as described below. Either of the described processes may be used to serve as an explicit SNMP provisioning command. Additionally, either process may be used to support dynamic allocation. The dynamic allocation may be used to support applications such as MPLS explicit routing, for example.

STS-1 level bandwidth dynamic-allocation:

Note: here the bandwidth requirement should be N times of STS-1.

PROCESS 2:

For initiating a CAP nodeA:

10

Step 1: Receive IP destination address.

Step2: Find corresponding NSAP address through Address Binding Table. Then based on Path Table, find the physical path to reach the next CAP.

15

Step3: Find next intermediate node's NSAP address based on information obtained from Path Table, Denote it as node B, Check the resource Header Table to find out the link which satisfies the following conditions:

The Neighbor of this link is B

The available bandwidth for this link exceeds the bandwidth as required

If there is no such link, Respond NO.

20

Step4: Send Convergence Protocol Bandwidth Allocation message to the next node, the message includes the source node NSAP address, the destination node NSAP address, required bandwidth, and identified STS-1 slots. Then wait for confirmation message from next node.

Step5: Get the Bandwidth Allocation Confirmation message from the next node.

25

If YES, modify corresponding resource Header Table and associated STS-1

Cao 3

20

Status Table to reflect the bandwidth allocation. Provision the corresponding framer. Respond YES.

If NO, respond NO.

For intermediate CIP node and destination CAP node:

5           Step 1: After receiving Bandwidth Allocation Message from previous node, check whether it is the final destination or not.

              If YES, respond YES through Bandwidth Allocation Confirmation message to previous node and provision corresponding STS-1s as path termination so that traffic inside this group of STS-1s will be processed by packet switching fabric, provision the corresponding framer.

10

Step2: Otherwise, it's not the final destination. Through path table, find out the NSAP address of the next, denote it as N.

Step3: Check the resource Header Table to find out the link which satisfies the following conditions:

15           The Neighbor of this link is N

              The available bandwidth for this link exceeds the bandwidth as required

              If there is no such link, Respond NO through Bandwidth Allocation Confirmation message.

20           Step4: Forward Convergence Protocol Bandwidth Allocation message to the next node, the message includes the source node NSAP address, the destination node NSAP address, required bandwidth, the STS-1 No. the current node allocated. Then wait for confirmation message from next node.

Step5: Get the confirmation message from next node.

              If YES, modify corresponding resource Header Table and associated STS-1

25           Status Table to reflect the bandwidth allocation. Setup corresponding cross-

Cao 3

21

connect between STS-1 slots as specified by previous node and STS-1 slots it allocated, record this information in cross-connect table and respond YES.

If NO, respond NO.

5

**PROCESS 3: STS-1 level bandwidth de-allocation:**

**For an initiating CAP node:**

**Step 1: Receiving IP destination address.**

10

**Step2: Find corresponding NSAP address of the given IP address through Address Binding Table.**

**Step3: Find next intermediate node's NSAP address based on information obtained from The Path Table. Find out STS1s allocated to reach the corresponding NSAP from the resource Header Table and associated STS-1 Status Table, choose exact No. of them based on the given requirement.**

15

**Step4: Send Convergence Protocol Bandwidth De-allocation message to the next node, the message includes the source node NSAP address, the destination node NSAP address, released bandwidth and those identified STS-1 slots.**

20

**Step5: Modify corresponding resource Header Table and associated STS-1 Status Table to reflect the bandwidth de-allocation, re-provision the framer to de-allocate this group of STS-1.**

**For an intermediate CIP node and destination CAP node:**

**Step 1: After receiving Bandwidth De-allocation Message from previous node, check whether it is the final destination or not.**

25

**If YES, re-provisioning the framer of those STS-1 as contained in the incoming message. Return.**

Cao 3

22

Step2: Otherwise, it's not the final destination. Through path table, find out the NSAP address of the next, denote it as N.

Step3: Based on cross-connect table, find out the corresponding STS-1 cross-entry.

5 Find the corresponding egress STS-1 slot entry and delete these cross-connect provisions.

Step4: Forward Convergence Protocol Bandwidth De-allocation message to the next node, the message includes the source node NSAP address, the destination node NSAP address, the required bandwidth and those identified STS-1 slot No. then return.

10 The topology discovery and dynamic bandwidth allocation just described may be used to establish an end-to-end multi-protocol label switched (MPLS) path that provides for minimal blocking. In order to support QoS enhanced MPLS, each CAP node should also support Label Distribution Protocol (LDP), as described in, L. Anderson et al, "LDP Specification, available at  
15 <http://www.ietf.org/internet-drafts/draft-ietf-mpls-ldp-03.txt>, January 1999, and Constrained Routing Label Distribution Protocol (CRLDP) as discussed in, B.Jamoussi, et al, "Constraint-based LSP Setup Using LDP,"  
20 <http://www.ietf.org/internet-drafts/draft-ietf-mpls-ldp-01.txt>, both of which are hereby incorporated by reference in their entirety. In this exemplary embodiment, each CAP acts as a Label Switching Router (LSR). If one assumes that the CAP is positioned as core (or intermediate) LSR, the responsibility of edge LSR converts the conventional IP header into a label and initiates the LDP Path setup message.

25 To support explicit routing based traffic engineering, a CRLDP message is used to explicitly setup a Label Switching Path. Based on information advertised by each CAP, an external router will have clear topology information about this virtual autonomous system. Consequently, the external router is capable of setting a path traversing this virtual autonomous system.



Cao 3

23

Beyond best-effort delivery, in order to support end-to-end QoS label switching, a system in accordance with the principles of the present invention sends traffic contract associated with the corresponding IP flow along the explicit path. As a result, each node along the path can decide in advance whether it can support or deny the traffic contract request.

The process used to set up end-to-end QoS based LSP follows. It is assumed that CRLDP is used to reserve bandwidth instead of RSVP, since CRLDP is based on hard-state implementation. Since RSVP is based on soft-state protocol implementation, periodic state-refreshing may consume formidable bandwidth and computing, and CRLDP is therefore preferred.

Regarding the control path, when an external core router receives the IP flow's traffic request, core router forwards this request to an adjacent CAP, denoting it as A, if next node, as indicated in the explicit path, is an adjacent CAP. Node A then generates the equivalent bandwidth needed to reach next CAP, based on the given traffic contract and the Connection Admission Control mechanism. Node A then looks up its resource table to determine whether existing ports can satisfy this request or not. If Yes, Node A reserves the bandwidth and forwards the request to next node. On the other hand, if Node A's existing ports cannot satisfy the request, Node A calculates the closest No. of STS-1 needed to setup a new port to satisfy this request, then determines the internal path (tributary) based on:

(1). Next Node (CAP) information contained in the given request

(2). Physical path to reach this CAP based on Address Binding Table and Path Table

(3) Based on Process 2 Node A either provisions the path to accept this flow

Cao 3

24

or denies this request. If this IP flow can be accommodated, the new Convergence Protocol-based Tunneled CRLDP message is sent to next node.

5 (4) If the next node is CIP (instead of destination CAP), after receiving Convergence Protocol based Tunneled CRLDP message, it will get the next node information based on path table, and forward it to next node, until it reaches the next CAP.

10 (5) The next CAP terminates the request, calculate its egress link, determines whether it can support it or not. If not, sends "No" back to entry CAP through Tunneled CRLDP message. Otherwise, forward request to next node (router/CAP), if any, and wait for the response.

15 (6) After the next CAP gets the response (assuming positive, otherwise, forward NO back to the entry CAP) which includes the egress label, allocates its own ingress label for this IP flow, forward it to entry CAP through Tunneled CRLDP message including the label it allocated.

(7) After source CAP gets the response which includes the egress label allocated by destination CAP, allocates its own ingress label for this IP flow, forward it to the external core router.

20 Through the above control-message exchange, a system in accordance with the principles of the present invention sets up the LSP for the incoming IP flow traversing transport core. After establishing the LSP, the data flow is as follows.

(1) The packet of this IP flow is sent to entry CAP (denoting it as A) from the core router.

25 (2) Based on label exact matching, node A finds the corresponding internal channel/tributary to be used, swaps the ingress label into the egress label, and, if the next node is a CIP, Node A forwards the packet to next node.

Cao 3

25

(3) An intermediate SONET node passes the packet through this internal channel via the cross-connect fabric. This internal channel is a SONET path based on a cross-connect fabric that is established via the new convergence protocol signalling.

5 (4) The destination CAP terminates the internal channel/tributary via label exact matching, swaps the ingress label into egress label and sends the flow to next LSR.

10 Using this approach provides a flexible, bandwidth-adaptive backbone. Additionally, minimal end-to-end latency is needed, since only the entry CAP and exit CAP involve layer 2 label swapping and related packet flow queuing. Additionally, in the intermediate CIPs, layer-1 pass-through gives the minimal (deterministic) traversing latency.

15 In accordance with the principles of the invention, a bandwidth-on-demand SONET/SDH transport infrastructure may be effected using the new convergence protocol as set forth below.

20 In one illustrative embodiment, one which employs a centralized resource management approach, a network management system triggers bandwidth allocation. During the path-setup period, it sends the following information to the initiating node (only): the destination address and required bandwidth. Then Convergence Protocol is used to exchange bandwidth information among nodes along the path. Processes 2 and 3, as described above, can be used to support this approach.

25 Step 1: NMS sends physical path provisioning information to ATM/IP side of CAP S1, which includes the following information: (assuming through GUI interface)

Management ATM/IP address of CAP S2

Cao 3

26

Bandwidth required: In this example, assuming OC-3C

Logical link layer provisioning information (In IP case, Frame-relay or PPP provisioning information) for both ends

Logical port provisioning information (for both ends)

5      ATM/IP address information (for both ends)

Routing information provisioning (for IP case, OSPF/RIP/BGP, for ATM case, OPSF/PNNI)

10      Step 2: ATM/IP side of CAP S1 finds corresponding management NSAP address of CAP S2 through looking up its address-binding table based on given ATM/IP management address information of CAP S2.

If corresponding management NSAP address of CAP S2 can't be found, negative response is sent back to NMS. Otherwise, proceeds to Step 3.

15      Step 3: ATM/IP side of CAP S1 forwards provisioning information to its SONET side, which includes bandwidth requirement information and management NSAP address of CAP S2

Step 4: SONET side of CAP S1 first determines whether inter-connect bandwidth inside CAP is sufficient to satisfy this requirement. If not, it sends negative response back to ATM/IP side of CAP S1, and then this negative response is further forwarded to NMS.

20      Step 5: If inter-connect bandwidth can accommodate this, based on established SONET topology, SONET side of CAP S1 finds out the path to reach SONET side of CAP S2, it also finds out its egress ports it can use to reach CAP S2. If it can't, it sends negative response back to ATM/IP side of CAP S1, and then this negative response is further forwarded to NMS.

25      Step 6. Then SONET side of CAP S1 determines that for any of identified egress port, whether egress port bandwidth is sufficient to satisfy this requirement. If not, it sends negative response back to ATM/IP side of CAP S1, and then this response is further forwarded to NMS.

Cao 3

27

5       Step 7: If its egress port bandwidth can accommodate this requirement, based on identified path information, SONET side of CAP S1 reserves certain STS-1 slots and initiates the signaling message flow and sends it to next node and waits for response from next node. The message includes path information and reserved STS-1 slots information.

(Assuming source/explicit routing is used)

10       Step 8: If response received from next node is positive, SONET side of CAP S1 provisions corresponding cross-connects, and sends physical provisioning information through proprietary signaling to ATM/IP side of CAP S1. Then ATM/IP side will provision corresponding tributary, including physical layer, logic link layer and ATM/IP layer, then initiates routing stack. Then positive response is sent back to NMS (Navis in this case).

15       If response received from next node is negative, SONET side of CAP S1 cancels STS-1 slot reservation, then sends negative response back to ATM/IP side of CAP S1, and then this negative response is further forwarded to NMS.

20       Using a distributed resource management approach, each CAP node has the up-to-date transport network topology and each CAP collects the performance monitoring information. Based on the above information, it makes a decision by itself as to whether a link's bandwidth should be adjusted. Each CAP also  
25       establishes the path based on routing table-lookup; makes the decision to re-adjust the bandwidth for each established tributary based on performance monitoring, and exchanges the bandwidth allocation/de-allocation information with other nodes along the path through Convergence Protocol Processes 2 and 3, as described above. Referring to Figure 12, the system 1200 includes a network management system 1202, first and second users 1204 and 1206 that are respectively connected to ATM/IP networks 1208 and 1210. Routers R1 through R4 connect users 1204 and 1206 through the ATM/IP networks 1208 and 1210 to CAPs 1212 and 1214, respectively, which provide access to the transport facilities of the SONET/SDH system 1216.

Cao 3

28

Such a bandwidth allocation process will be described below with reference to the scenario diagram of Figure 13 in which:

Step 1: NMS sends physical path provisioning information to ATM/IP side of CAP1, which includes the following information: (assuming through GUI interface)

5           Management ATM/IP address of CAP2

Bandwidth required: in this example, assuming OC-3C

Logical link layer provisioning information (in IP case, Frame-relay or PPP provisioning information) for both ends

Logical port provisioning information (for both ends)

10          ATM/IP address information (for both ends)

Routing information provisioning (for IP case, OSPF/RIP/BGP, for ATM case, OPSF/PNNI)

15          Step 2: ATM/IP side of CAP1 finds corresponding management NSAP address of CAP S2 through looking up its address-binding table based on given ATM/IP management address information of CAP2.

If corresponding management NSAP address of CAP2 can't be found, negative response is sent back to NMS. Otherwise, proceeds to Step 3.

20          Step 3: ATM/IP side of CAP1 forwards provisioning information to its SONET side, which includes bandwidth requirement information and management NSAP address of CAP2

Step 4: SONET side of CAP1 first determines whether inter-connect bandwidth inside CAP is sufficient to satisfy this requirement. If not, it sends negative response back to ATM/IP side of CAP1, and then this negative response is further forwarded to NMS.

25          Step 5: If inter-connect bandwidth can accommodate this, based on established SONET topology, SONET side of CAP1 finds out the path to reach SONET side of CAP2, it also finds out its egress ports it can use to reach CAP2. If

Cao 3

29

it can't, it sends negative response back to ATM/IP side of CAP1, and then this negative response is further forwarded to NMS.

5 Step 6: Then SONET side of CAP1 determines that for any of identified egress port, whether egress port bandwidth is sufficient to satisfy this requirement. If not, it sends negative response back to ATM/IP side of CAP1, and then this response is further forwarded to NMS.

10 Step 7: If its egress port bandwidth can accommodate this requirement, based on identified path information, SONET side of CAP1 reserves certain STS-1 slots and initiates the signaling message flow and sends it to next node and waits for response from next node. The message includes path information and reserved STS-1 slots information.

(Assuming source/explicit routing is used)

15 Step 8: If response received from next node is positive, SONET side of CAP1 provisions corresponding cross-connects, and sends physical provisioning information through proprietary signaling to ATM/IP side of CAP1. Then ATM/IP side will provision corresponding tributary, including physical layer, logic link layer and ATM/IP layer, then initiates routing stack. Then positive response is sent back to NMS (Navis in this case).

20 If response received from next node is negative, SONET side of CAP1 cancels STS-1 slot reservation, then sends negative response back to ATM/IP side of CAP1, and then this negative response is further forwarded to NMS in steps 9 through 11.

25 In another illustrative embodiment in accordance with the principles of the present invention, network employs protocol-driven resource management. With this approach, a bandwidth reservation protocol such as CRLDP or RSVP etc. may be used to trigger bandwidth allocation. As previously described, these protocols embody implicit bandwidth requirement (to support the corresponding QoS) and explicit path information. Based on a Connection Admission Control algorithm, the entry CAP node will convert the implicit bandwidth requirement into  
30 equivalent bandwidth, then exchange bandwidth allocation/de-allocation

Cao 3

30

information with other nodes along the path as indicated in the incoming protocol message.

5 The flowchart of Figure 14 illustrates the process of initializing, establishing a CAP path and dynamically allocating bandwidth in accordance with the principles of the present invention. The de-allocation of bandwidth in accordance with the principles of the present invention is also illustrated in the flowchart. In step 1400 the process begins and proceeds to step 1402 where nodes employing the convergence protocol are initialized. The initialization includes the running of both OSI stack and TCP/IP stack in CAPs and the running of OSI stack in CIPs.

10 The process proceeds from step 1402 to step 1404 where network nodes follow a process of auto-discovery as previously described in greater detail in the above section entitled, "IS-IS based topology auto-discovery". Following step 1404 the process proceeds to step 1406 where each CAP pair establishes a logical path between themselves. This logical path establishment may entail the flooding of a CAP's IP address information to other CAPs employing a label switched path variable length field. The logical path establishment process is described in greater detail in the above section entitled "CAP path establishment".

15

From step 1406 the process proceeds to step 1408 where a receiving node responds to the reception of a IS-IS label switched path message in accordance with the principles of the present invention by treating it in a conventional manner if the receiving node is a CIP. A CAP will, additionally, refresh its routing table, based on the incoming information. This process is discussed in greater detail in relation to Figure 7. From step 1408 the process proceeds to step 1410 where a CAP establishes an address binding table and obtains the path information necessary to reach other CAPs. This process is discussed in greater detail in relation to "Process 1", described above. From step 1410 the process proceeds to step 1412 where a node passes a convergence protocol bandwidth allocation message node to node until it reaches the destination CAP. In step 1414 the bandwidth along the path is allocated, if available. The processes of steps 1412 and 1414 are described in greater detail in the discussion of "Process 2",

20

25

30



Cao 3

31

described above. At some point there may be a need for bandwidth de-allocation and, in that case, the process would proceed to step 1416 where a node passes a convergence protocol bandwidth de-allocation message, node to node, until the destination CAP is reached. In step 1418 the bandwidth is de-allocated, as  
5 appropriate. The processes of steps 1416 and 1418 are described in greater detail in the discussion of "Process 3", described above. The overall process may proceed to END in step 1420, for example, during maintenance or installation operations, for example.

The foregoing description of specific embodiments of the invention has  
10 been presented for the purposes of illustration and description. It is not intended to be exhaustive or to limit the invention to the precise forms disclosed, and many modifications and variations are possible in light of the above teachings. The embodiments were chosen and described to best explain the principles of the invention and its practical application, and to thereby  
15 enable others skilled in the art to best utilize the invention. It is intended that the scope of the invention be limited only by the claims appended hereto. What is claimed is:

#### 4. Brief Description of Drawings

Written above.

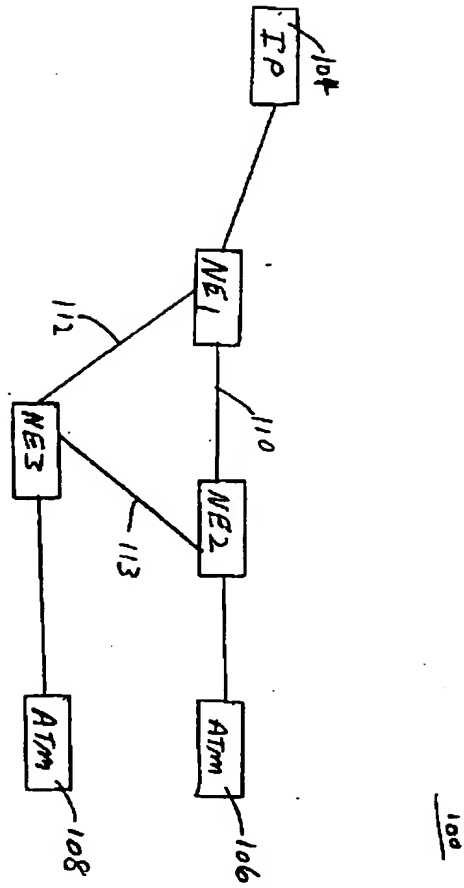


FIGURE 1

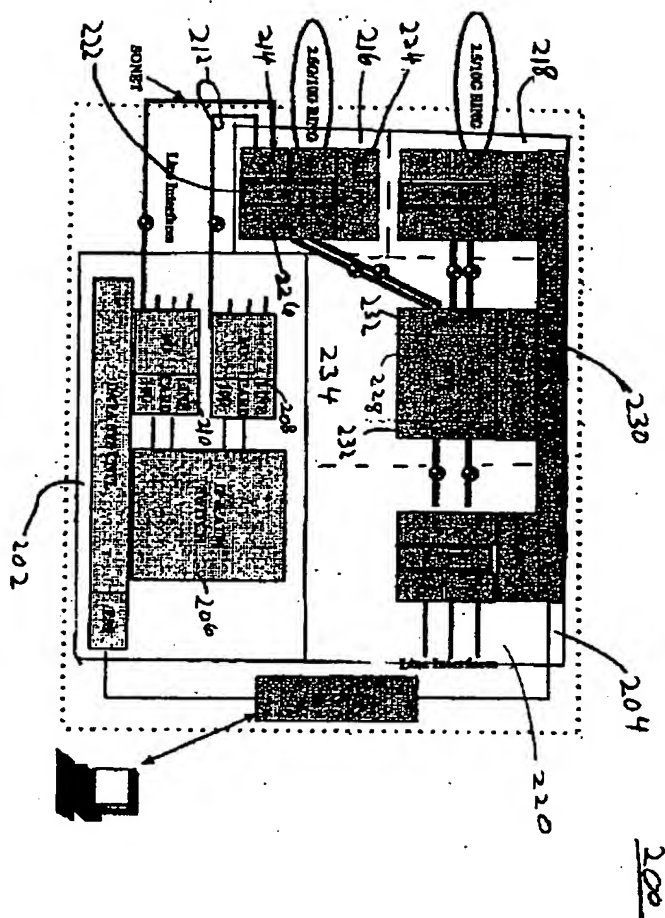


FIGURE 2

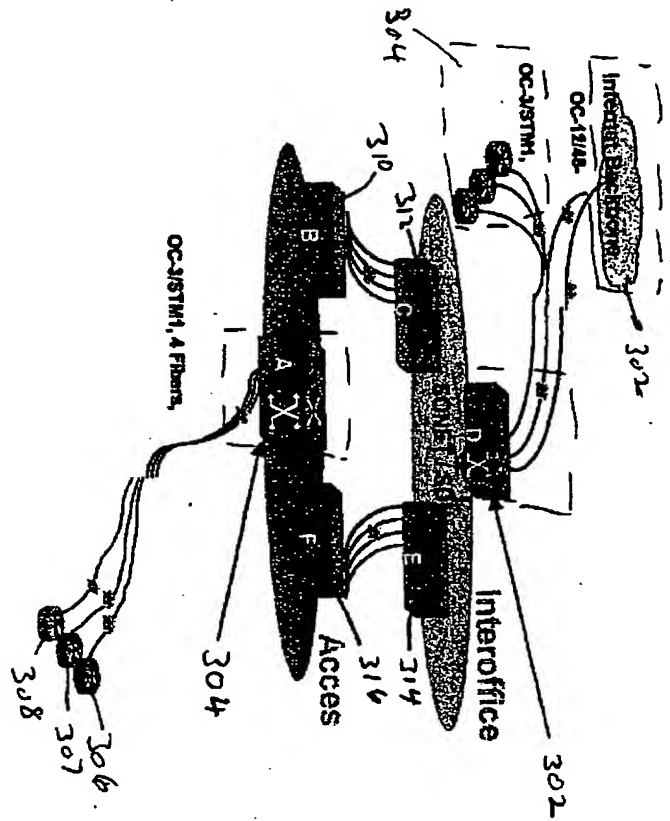


FIGURE 3

Byte 1	Convergence Protocol Discriminator	
Byte 2	Length Indicator	
Byte 3	Length indicator (4 bits)	Version (4 bits)
Byte 4	PDU Type	
Byte 5	Checksum	
	Variable Length Fields	

Figure 4

Byte 1	Intradomain Routing Protocol Discriminator			
Byte 2	Length Indicator			
Byte 3	Version/Protocol ID Extension			
Byte 4	ID Length			
Byte 5	R	R	R	PDU Type
Byte 6	Version			
Byte 7	Reserved			
Byte 8	Maximum Area Addresses			
Byte 9	PDU Length			
Byte 10	PDU Length			
Byte 11	Remaining lifetime			
Byte 12	Remaining lifetime			
Byte 13	LSP ID			
Byte 14	Sequence number			
Byte 15	Sequence number			
Byte 16	Sequence number			
Byte 17	Sequence number			
Byte 18	Checksum			
Byte 19	Checksum			
Byte 20	P	ATT	LSPDBOL	IS Type
	Variable Length Fields			

Figure 5

Byte 1	CODE
Byte 2	LENGTH
	Value

Figure 6

Byte 1	CODE = 15
Byte 2	LENGTH = 4
	IP-V4 Address

Byte 1	CODE = 16
Byte 2	LENGTH = 16
	IP-V6 Address

Figure 7

IP Address (4 bytes if IP-V4 or 7 bytes if IP-V6 is supported)	Corresponding NSAP Address (20 bytes)
---	--

Figure 8

Source Node's NSAP Address	Destination Node's NSAP Address	Next Node's NSAP Address
----------------------------	---------------------------------	--------------------------

Figure 9

Neighbor NSAP Address (20 bytes)	Available Bandwidth (16 bytes: unit: STS1)	STS-1 Array Pointer (4 bytes – 8 bytes, depending on 32-bit mode or 64-bit mode)
-------------------------------------	---	---

Figure 10

STS-1 Slot No.	Free Status	Destination NSAP Address	Available Bandwidth
----------------	-------------	-----------------------------	------------------------

Figure 11



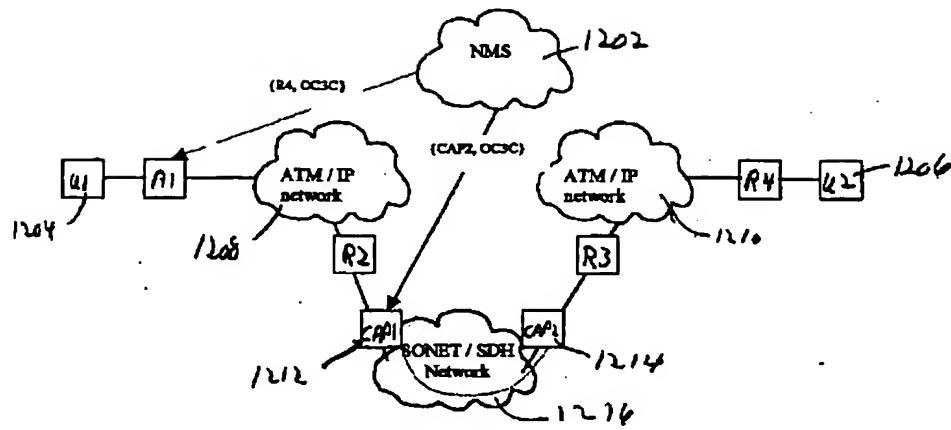
1200

Figure 12

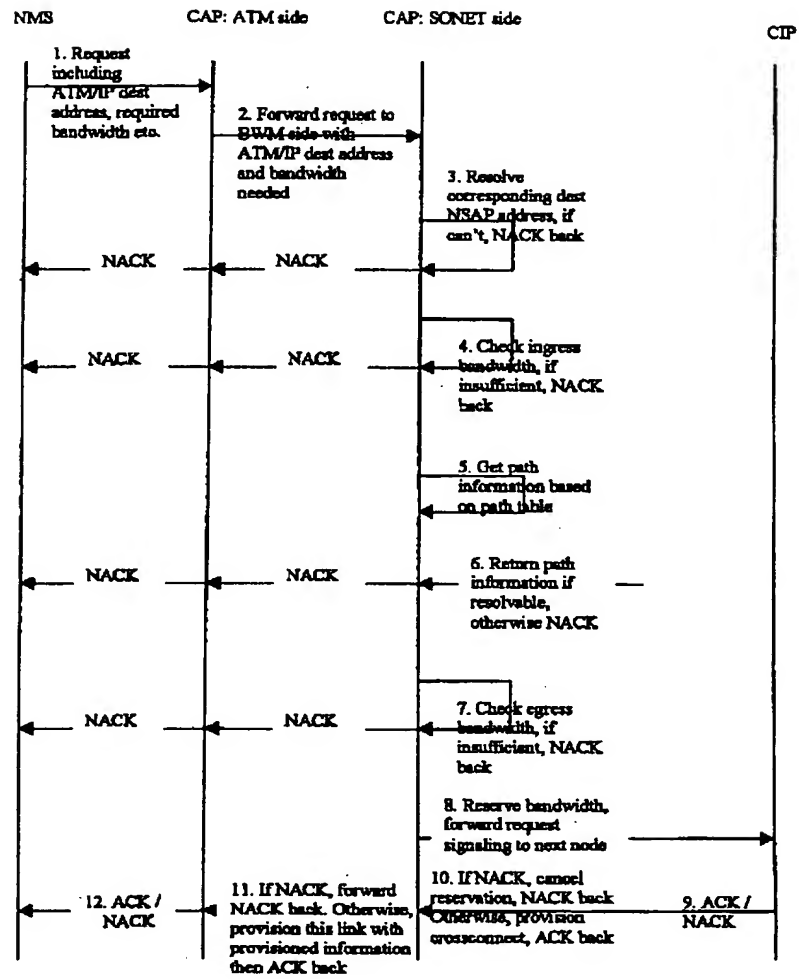
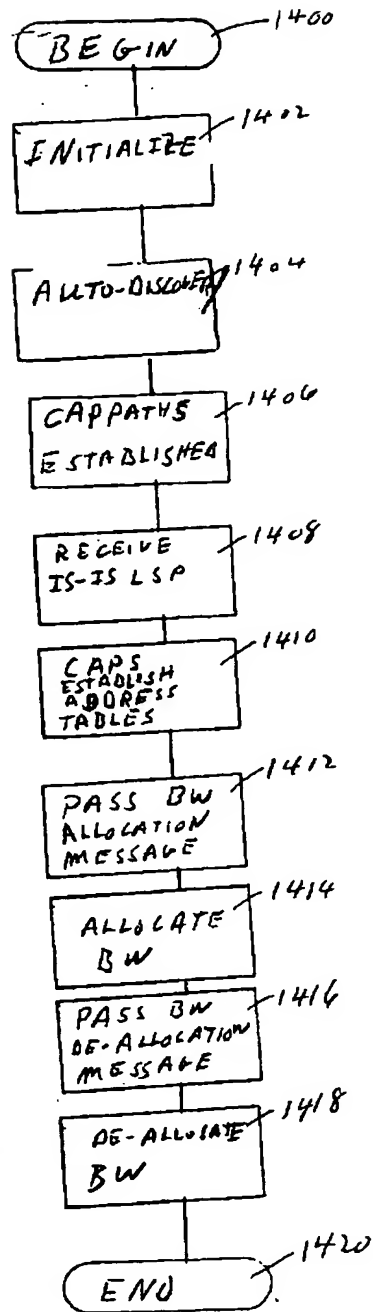


Figure 13

FIGURE 14



1 Abstract

A telecommunications system includes a convergence protocol that provides direct inter-layer communications between nodes within the network. The transport layer of one network node may communicate directly with the service layer of another network node. The service layer may be configured as an internet protocol (IP) layer, which may employ asynchronous transfer mode switching, and the transport layer may be a SONET transport layer. Each of the nodes that runs the convergence protocol may be a telecommunications element such as an add-drop-multiplexer (ADM) or a digital cross connect. SONET bandwidth provisioning may be accomplished through the inter-element, inter-layer communications of the convergence protocol.

## 2 Representative Drawing

### Figure 1